

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
19 July 2001 (19.07.2001)

PCT

(10) International Publication Number  
WO 01/51627 A2

- (51) International Patent Classification: C12N 15/00 (74) Agent: MARSH, David, R.; Howrey Simon Arnold & White, LLP, 1299 Pennsylvania Avenue, N.W., Box 34, Washington, DC 20004-2402 (US).
- (21) International Application Number: PCT/US01/00552
- (22) International Filing Date: 5 January 2001 (05.01.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 60/174,880 7 January 2000 (07.01.2000) US
- (71) Applicant: MONSANTO COMPANY [US/US]; 800 N. Lindbergh Boulevard, St. Louis, MO 63167 (US).
- (72) Inventors: HAUGE, Brian, M.; 17 Clifton Avenue, Beverly, MA 01915 (US). WANG, Ming, Li; 312 Lowell Street, Lexington, MA 02420 (US). PARSONS, Jeremy, D.; 37 Walpole Road, Cambridge CB1 3TH (GB). PAR-NELL, Laurence, D.; 9 Eustis Street #2L, Cambridge, MA 02140-2226 (US).
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- Published: — without international search report and to be republished upon receipt of that report

[Continued on next page]

(54) Title: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH SOYBEAN CYST NEMATODE RESISTANCE

rhg1 LRR

consensus LRR	L F S N L P	N L E E L D L S N N	L T	S L P P G
	a x x a x	x L x x L x L x x N	L	I
LRR 177-200	T L G L L P	G L R K L S L H D N Q	I G	G S I P S
LRR 201-224	S L G F C P	N L R G V Q L F N N R	L T	G S I P L
LRR 225-248	L A N S T	L L Q S L D L S N N L	L T	G A I P Y S
LRR 249-272	L T H S F	K L Y W L N L S F N S	F S	G P L P A S
LRR 273-297	W G G N S K	S L T F L S L Q N N N	L S	G S L P N S
LRR 302-325	L G S L R	R L Q N L I L D H N F	F T	G D V P A S
LRR 326-349	I G T L S	E L N E L S L S H N K	F S	G A I P N E
LRR 350-373	T L S N L S	R L K T L D I S N N A	L N	G N L P A
LRR 374-397	S L G R L R	S L T L L N A E N N L	L D	N Q I P Q
LRR 398-421	S I A N I S	N L S V L I L S R N Q	F S	G H I P S
LRR 422-445	S F I D S Q R	S L R Q L D L S L N N	F S	G E I P V
LLR 446-470	L L A K K F N S	L N L F N V S N S	L S	G S V P P

(57) Abstract: The present invention is in the field of soybean genetics. More specifically, the invention relates to nucleic acid molecules from regions of the soybean genome, which are associated with soybean cyst nematode resistance. The invention also relates to proteins encoded by such nucleic acid molecules as well as antibodies capable of recognizing these proteins. The invention also relates to nucleic acid markers from regions of the soybean genome, which are associated with soybean cyst nematode resistance. Moreover, the invention relates to uses of such molecules, including, transforming soybean cyst nematode resistant soybean with constructs containing nucleic acid molecules from regions of the soybean genome, which are associated with soybean cyst nematode resistance. Furthermore, the invention relates to the use of such molecules in a plant breeding program.

WO 01/51627 A2



---

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

## NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH SOYBEAN CYST NEMATODE RESISTANCE

### FIELD OF THE INVENTION

The present invention is in the field of soybean genetics. More specifically, the invention relates to nucleic acid molecules from regions of the soybean genome, which are associated with soybean cyst nematode (SCN) resistance. The invention also relates to proteins encoded by such nucleic acid molecules as well as antibodies capable of recognizing these proteins. The invention also relates to nucleic acid markers from regions of the soybean genome, which are associated with SCN resistance. Moreover, the invention relates to uses of such molecules, including, transforming SCN sensitive soybean with constructs containing nucleic acid molecules from regions in the soybean genome, which are associated with SCN resistance. Furthermore, the invention relates to the use of such molecules in a plant breeding program.

### BACKGROUND OF THE INVENTION

The soybean, *Glycine max* (L.) Merrill (*Glycine max* or soybean), is one of the major economic crops grown worldwide as a primary source of vegetable oil and protein (Sinclair *et al.*, *Compendium of Soybean Diseases*, 3<sup>rd</sup> Ed. APS Press, St. Paul, MN, p. 106. (1989)). The growing demand for low cholesterol and high fiber diets has also increased soybean's importance as a health food.

Prior to 1940, soybean cultivars were either direct releases of introductions brought from Asia or pure line selections from genetically diverse plant introductions. The soybean plant was primarily used as a hay crop in the early part of the 19th century. Only a few introductions were large-seeded types useful for feed grain and oil production. From the mid 1930's through the 1960's, gains in soybean seed yields were achieved by changing the breeding method from evaluation and selection of introduced germplasm to crossing elite by elite lines. The continuous cycle of cross hybridizing the elite strains selected from the progenies of previous crosses resulted in the modern day cultivars.

Over 10,000 soybean strains have now been introduced into the United States since the early 1900's (Bernard *et al.*, *United States National Germplasm Collections*. In: L.D. Hil (ed.), *World Soybean Research*, pp. 286-289. Interstate Printers and Publ., Danville, IL (1976)). A limited number of those introductions form the genetic base of cultivars developed from the hybridization and selection programs (Johnson *et al.*, *The Soybean*, Norman Ed., Academic Press, N.Y., pp. 1-73 (1963)). For example, in a survey conducted by Specht and Williams, *Genetic Contributions*, Fehr eds. American Soil Association, Wisconsin, pp. 49-73 (1984), for the 136 cultivars released from 1939 to 1989, only 16 different introductions were the source of cytoplasm for 121 of that 136. Certain soybean strains are sensitive to one or more pathogens. One economically important pathogen is SCN.

SCN accounts for roughly 40% of the total disease in soybean and can result in significant yield losses (up to 90%). SCN is the most destructive pest of soybean to date and accounts for an estimated yield loss of up to \$809 million dollars annually. Currently, the most cost effective control measures are crop rotation and the use of host plant resistance. While breeders have successfully developed SCN resistant soybean lines, breeding is both difficult and time consuming due to the complex and polygenic nature of resistance. The resistance is often race specific and does not provide stability over time due to changing SCN populations in the field. In addition, many of the resistant soybean varieties carry a significant yield penalty when grown in the absence of SCN.

SCN, *Heterodera glycines Ichinohe*, was identified on soybeans in the United States in 1954 at Castle Hayne, N.C. Winstead, *et al.*, *Plant Dis. Rep.* 39:9-11 (1955). Since its discovery the SCN has been recognized as one of the most destructive pests in soybean. It has been reported in nearly all states in which soybeans are grown, and it causes major production problems in several states, being particularly destructive in the Midwestern states. See generally: Caldwell, *et al.*, *Agron. J.* 52:635-636 (1960); Rao-Arelli *et al.*, *Crop. Sci.* 28:650-652, (1988); Baltazar *et al.*, *Soybean Genet. Newsl.* 19:120-122 (1992); Concibido, *et al.*, *Crop. Sci.*, (1993). For example, sensitive soybean

cultivars had 5.7-35.8% lower seed yields than did resistant cultivars on SCN race-3 infested sites in Iowa. (Niblack *et al.*, *Plant Dis.* 76:943-948 (1992)).

Shortly after the discovery of SCN in the United States, sources of SCN resistance were identified (Ross *et al.*, *Plant Dis. Rep.* 41:923-924 (1957)). Some lines such as Peking and Plant Introduction (PI) PI88788, were quickly incorporated into breeding programs. Peking became widely used as a source of resistance due to its lack of agronomically undesirable traits, with Pickett as the first SCN resistant cultivar released (Brim *et al.*, *Crop Sci.* 6:305 (1966)). The recognition that certain SCN resistant populations could overcome resistant cultivars lead to an extensive screen for additional sources of SCN resistance. PI88788 emerged as a popular source of race 3 and 4 resistance even though it had a cyst index greater than 10% (but less than 20%) against race 4, and Peking and its derivatives emerged as a popular source for races 1 and 3. PI437654 was subsequently identified as having resistance to all known races and its SCN resistance was backcrossed into Forrest. Currently there are more than 130 PIs known to have SCN resistance.

SCN race 3 is considered to be the prominent race in the Midwestern soybean producing states. Considerable effort has been devoted to the genetics and breeding for resistance to race 3. While both Peking and PI88788 are resistant to SCN race 3, classical genetics studies suggest that they harbor different genes for race 3 resistance (Rao-Arelli *et al.*, *Crop Sci.* 28:650-652 (1988)). Crosses between PI88788(R) and Essex(S) segregate 9(R): 55(S) in the F<sub>2</sub> population and 1(R): 26(Seg): 37(S) families in the F<sub>3</sub> generation, suggesting that resistance to race 3 in PI88788 is conditioned by one recessive and two dominant genes, where as Peking and PI90763 resistance is conditioned by one dominant and two recessive genes. Based on reciprocal crosses, Peking, Forrest, and PI90763 have genes in common for resistance to SCN race 3 (Rao-Arelli *et al.*, *Crop Sci.*, 28:650-652 (1988)). A cross between Peking and PI88788 segregates 13(R):3(S) in the F<sub>2</sub> generation, indicating a major difference between the parents for race 3 resistance. Generation mean analysis based on four crosses between resistant and sensitive genotypes; A20 (R), Jack (R), Cordell (R) and A2234 (S), suggests that an additive genetic model is

sufficient to explain most of the genetic variation of race 3 SCN resistance in each cross, while the analysis of the pooled data indicates the presence of dominant effects as well (Mansur *et al.*, *Crop Sci.* 33:1249-1253 (1993)). This analysis further indicates that race 3 resistance is probably under the genetic control of three, but not more than four genes.

RFLP analysis of segregating populations between resistant and sensitive lines; PI209332 (R), PI90763 (R), PI88788 (R), Peking (R) and Evan (S), identified a major SCN resistance QTL (*rhg1*) which maps to linkage group G (Concibido *et al.*, *Theor Appl. Genet.* 93:234-241 (1996)). In this study, *rhg1* explains 51.4% of the phenotypic variation in PI209332, 52.7% of the variation in PI90763, 40.0% of the variation in PI88788 and 28.1% of the variation in Peking. This major resistance QTL was assumed be one and the same in all of the mapping populations employed. However, as pointed out by the authors, it is possible that the genomic interval contains distinct but tightly linked QTLs. In a related study using PI209332 as the source of resistance, Concibido *et al.*, *Crop Sci.* 36:1643-1650 (1996), show that a QTL on linkage group G (*rhg1*) is effective against the three SCN races tested, explaining 35% of the phenotypic variation to race 1, 50% of the variation to race 3, and 54% of the variation to race 6. In addition to the major QTL on linkage group G, 4 other QTLs mapping to linkage groups D, J, L and K were identified, with some of the resistance loci behaving in a race specific manner.

Concibido *et al.* (*Crop Sci.* 37:258-264 (1997)) found significant association of marker C006V to a major QTL on linkage group G (*rhg1*) and resistance to race 1, race 3 and race 6, in Peking and PI90763 (Evan X Peking, Evan X PI90763) and races 3 and 6 in PI88788 (Evan X PI88788), in agreement with the previous study based on the P209332 source of resistance (Concibido *et al.*, *Crop Sci.* 36:1643-1650 (1996)). The resistance locus near C006V was effective against all races tested in all of the resistance sources. While statistically significant against all races, this locus accounts for different proportions of the total phenotypic variation with the races tested. For example, in PI90763 the resistance locus near C006V explains more than three times the phenotypic variation against race 1 than against race 3. The variability can be attributed to

differences in the genetic backgrounds, variability among the SCN populations or may be a reflection of the limited size of the plant populations which were employed. This study further identified three additional independent SCN resistance QTLs; one near the RFLP marker A378H mapping to the opposite end of linkage group G from C006V (*rhg1*), one near the marker B032V-1 on linkage group J and a third linked to A280Hae-1 on linkage group N. Comparisons between the different SCN races indicated that some of the putative SCN QTLs behave in a race specific manner.

PI437654 was identified as having resistance to all known races. Based on analysis of 328 recombinant inbred lines (RIL) derived from a cross between PI437654 and BSR101, Webb reported six QTLs associated with SCN resistance on linkage groups A2, C1, G, M, L25 and L26 (U.S. Patent 5,491,081). An allele on linkage group G, presumed to be *rhg1*, is involved with certain SCN races tested (races 1, 2, 3, 5 and 14), and has the largest reported phenotypic effect on resistance to every race. In contrast, the QTLs on linkage groups A2, C1, M, L25 and L26 act in a race specific manner. The QTL on linkage group L25 was reportedly involved with four of the five races, while the QTLs on linkage groups, A2, C1 and L26 were each involved in resistance to two of the five races (U.S. Patent 5,491,081). Webb further reports data that the resistance to any of the five races is likely to result from the combined effects of the QTL involved in each race (U.S. Patent 5,491,081).

Qui *et al.* (*Theor Appl Genet* 98:356-364 (1999)) screened 200 F<sub>2:3</sub> families derived from a cross between Peking and Essex and identified RFLP markers which are associated with SCN resistance QTLs on linkage groups B, E, I and H. The three QTLs on linkage groups B, E and H jointly account for 57.7% of the phenotypic variation to race 1, the QTLs on linkage groups H and B account for 21.4% of the variation to race 3, while the QTLs on linkage groups I and E are associated with resistance to race 5 accounting for 14.0% of the phenotypic variation. In contrast to previous mapping studies which use Peking as the source of resistance, no significant association was detected to the *rhg1* locus on linkage group G. The authors point out that the marker Bng122, which

has been shown to have significant linkage to *rhg1*, is not polymorphic in the population employed (Concibido *et al.*, *Crop Sci.* 36:1643-1650 (1996)).

It has been reported that the *rhg1* locus on linkage group G is necessary for the development of resistance to any of the SCN races. There have been efforts to develop molecular markers to identify breeding lines harboring the *rhg1* SCN resistant allele. One of the most commonly used markers for marker assisted selection (MAS) of *rhg1* is an SSR locus that co-segregates and maps roughly 0.4 cM from *rhg1*. This SRR marker, BARC-Satt\_309 is able to distinguish most, if not all, of the SCN sensitive genotypes from those harboring *rhg1* from important sources of resistance such as Peking and PI437654. Two simple sequence repeat markers have been reported that can be used to select for SCN resistance at the *rhg1* locus (Concibido *et al.*, *Theor Appl Genet* 99: 811-818 (1999)). Satt\_309 was also effective in distinguishing SCN resistant sources PI88788 and PI209332 in many, but not all, sensitive genotypes. In particular, Satt\_309 can not be used for MAS in populations developed from "typical" southern US cultivars (*e.g.*, Lee, Bragg and Essex) crossed with resistance sources PI88788 or PI209332.

Matson *et al.* have reported a dominant SCN resistance locus, *Rhg4*, which is tightly linked to the '*i*' locus on linkage group A2 (Matson *et al.*, *Crop Sci.* 5:447 (1965)). The QTL reported by Webb on linkage group A2 maps near the '*i*' locus and is considered to be *Rhg4* (U.S. Patent 5,491,081). Webb concludes that only two loci on linkage groups A2 (*Rhg4*) and G (*rhg1*) explain the genetic variation to race 3.

#### SUMMARY OF THE INVENTION

The present invention includes and provides a method for the production of a soybean plant having an *rhg1* SCN resistant allele comprising: (A) crossing a first soybean plant having an *rhg1* SCN resistant allele with a second soybean plant having an *rhg1* SCN sensitive allele to produce a segregating population; (B) screening the segregating population for a member having an *rhg1* SCN resistant allele with a first nucleic acid molecule capable of specifically hybridizing to linkage group G, wherein the first



nucleic acid molecule specifically hybridizes to a second nucleic acid molecule that is linked to the *rhg1* SCN resistant allele; and, (C) selecting the member for further crossing and selection.

The present invention includes and provides a method of investigating an *rhg1* haplotype of a soybean plant comprising: (A) isolating nucleic acid molecules from the soybean plant; (B) determining the nucleic acid sequence of an *rhg1* allele or part thereof; and, (C) comparing the nucleic acid sequence of the *rhg1* allele or part thereof to a reference nucleic acid sequence.

The present invention includes and provides a method of introgressing SCN resistance or partial SCN resistance into a soybean plant comprising: performing marker assisted selection of the soybean plant with a nucleic acid marker, wherein the nucleic acid marker specifically hybridizes with a nucleic acid molecule having a first nucleic acid sequence that is physically linked to a second nucleic acid sequence that is located on linkage group G of soybean A3244, wherein the second nucleic acid sequence is within 500 kb of a third nucleic acid sequence which is capable of specifically hybridizing with the nucleic acid sequence of SEQ ID NO: 5, 6, complements thereof, or fragments thereof having at least 15 nucleotides; and, selecting the soybean plant based on the marker assisted selection.

The present invention includes and provides a method for the production of a soybean plant having an *Rhg4* SCN resistant allele comprising: (A) crossing a first soybean plant having an *Rhg4* SCN resistant allele with a second soybean plant having an *Rhg4* SCN sensitive allele to produce a segregating population; (B) screening the segregating population for a member having an *Rhg4* SCN resistant allele with a first nucleic acid molecule capable of specifically hybridizing to linkage group A2, wherein the first nucleic acid molecule specifically hybridizes to a second nucleic acid molecule linked to the *Rhg4* SCN resistant allele; and, (C) selecting the member for further crossing and selection.

The present invention includes and provides a method of investigating an *Rhg4* haplotype of a soybean plant comprising: (A) isolating nucleic acid molecules from the soybean plant; (B) determining the nucleic acid sequence of an *Rhg4* allele or part thereof; and (C) comparing the nucleic acid sequence of the *Rhg4* allele or part thereof to a reference nucleic acid sequence.

The present invention includes and provides a method of introgressing SCN resistance or partial SCN resistance into a soybean plant comprising: performing marker assisted selection of the soybean plant with a nucleic acid marker, wherein the nucleic acid marker specifically hybridizes with a nucleic acid molecule having a first nucleic acid sequence that is physically linked to a second nucleic acid sequence that is located on linkage group A2 of soybean A3244, wherein the second nucleic acid sequence is within 500 kb of a third nucleic acid sequence which specifically hybridizes with the nucleic acid sequence of SEQ ID NO: 7, complements thereof, or fragments thereof having at least 15 nucleotides; and, selecting the soybean plant based on the marker assisted selection.

The present invention includes and provides a substantially purified nucleic acid molecule comprising a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 5, 6, 8-23, 28-43, complements thereof, and fragments of either.

The present invention includes and provides a substantially purified first nucleic acid molecule with nucleic acid sequence which specifically hybridizes to a second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NOs: 5, 6, 8-23, 28-43.

The present invention includes and provides a substantially purified nucleic acid molecule comprising a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 7, 44-47, and 50-53, complements thereof, and fragments of either.

The present invention includes and provides a substantially purified first nucleic acid molecule with nucleic acid sequence which specifically hybridizes to a second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NOs: 50-53.

The present invention includes and provides a substantially purified protein or fragment thereof comprising an amino acid sequence selected from the group consisting of SEQ ID NOs: 1097, 1098, and 1100-1115 and fragments thereof.

The present invention includes and provides a substantially purified protein or fragment thereof comprising an amino acid sequence selected from the group consisting of SEQ ID NOs 1099, and 1116-1119 and fragments thereof.

The present invention includes and provides a transformed plant having a nucleic acid molecule which comprises: (A) an exogenous promoter region which functions in a plant cell to cause the production of a mRNA molecule; (B) a structural nucleic acid molecule encoding a protein or fragment thereof comprising an amino acid sequence selected from the group consisting of SEQ ID NOs: 1097, 1100, 1098, 1101, 1102-1115; and (C) a 3' non-translated sequence that functions in the plant cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule.

The present invention includes and provides a transformed plant having a nucleic acid molecule which comprises: (A) an exogenous promoter region which functions in a plant cell to cause the production of a mRNA molecule; (B) a structural nucleic acid molecule encoding a protein or fragment thereof comprising an amino acid sequence selected from the group consisting of SEQ ID NOs: 1099, 1116-1119; and (C) a 3' non-translated sequence that functions in the plant cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule.

The present invention includes and provides a transgenic seed having a nucleic acid molecule which comprises: (A) an exogenous promoter region which functions to cause the production of a mRNA molecule; (B) a structural nucleic acid molecule encoding a protein or fragment thereof comprising an amino acid sequence selected from the group consisting of SEQ ID NOs: 1097, 1100, 1098, 1101, 1102-1115; and (C) a 3' non-translated sequence that functions to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule.

The present invention includes and provides a transgenic seed having a nucleic acid molecule which comprises: (A) an exogenous promoter region which functions to

cause the production of a mRNA molecule; (B) a structural nucleic acid molecule encoding a protein or fragment thereof comprising an amino acid sequence selected from the group consisting of SEQ ID NOs: 1099, 1116-1119; and (C) a 3' non-translated sequence that functions to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule.

#### DESCRIPTION OF THE FIGURES

Figure 1 is an amino acid sequence alignment of the leucine rich repeat domain of *rhg1*.

Figure 2 is an amino acid sequence alignment of the leucine rich repeat domain of *Rhg4*.

#### DESCRIPTION OF THE SEQUENCE LISTINGS

The following sequence listings form part of the present specification and are included to further demonstrate certain aspects of the present invention. The invention may be better understood by reference to one or more of these sequences in combination with the detailed description presented herein.

SEQ ID NOs: 1-7 and 1097-1099 all refer to sequences from the line A3244.

SEQ ID NO: 1 is sequence ID 515O02\_region\_G2 from line A3244, and is adjacent to the contig containing *rhg1*.

SEQ ID NO: 2 is sequence ID 240O17\_region\_G3 from line A3244, and contains the *rhg1*, v.1 four exon gene at coding coordinates 45163-45314, 45450-45509, 46941-48763, 48975-49573. The amino acid translation for SEQ ID NO: 2 is SEQ ID NO: 1097.

SEQ ID NO: 3 is sequence ID 240O17\_region\_G3 from line A3244, and contains the *rhg1*, v.2 two exon gene at coding coordinates 46798-48763 and 48975-49573. The amino acid translation for SEQ ID NO: 3 is SEQ ID NO: 1098.

SEQ ID NO: 4 is sequence ID 318O13\_region\_A3 from line A3244, contains the *Rhg4* gene at coding coordinates 111805-113968 and 114684-115204, and has an amino acid translation of SEQ ID NO: 1099.

SEQ ID NO: 5 is sequence ID 240O17\_region\_G3\_8\_mRNA, and comprises the two *rhg1*, *v.2* exons from the coding sequence portion of SEQ ID NO: 3.

SEQ ID NO: 6 is sequence ID 240O17\_region\_G3\_8\_cds, and comprises the four *rhg1*, *v.1* exons from the coding sequence portion of SEQ ID NO: 2.

SEQ ID NO: 7 is sequence ID 318O13\_region\_A3\_17\_cds, and comprises the *Rhg4* coding sequence portion from SEQ ID NO: 4.

SEQ ID NOs: 8-43 and 1100-1115 all refer to *rhg1* sequences.

SEQ ID NO: 8 is sequence ID *rhg1\_A3244\_amplicon* from line A3244, contains four *rhg1*, *v.1* exons at coding coordinates 113-264, 400-459, 1891-3713, and 3925-4523, and has an amino acid translation of SEQ ID NO: 1100 and 1097.

SEQ ID NO: 9 is sequence ID *rhg1\_A3244\_amplicon*, contains two *rhg1*, *v.2* exons at coding coordinates 1748-3713 and 3925-4523 and has an amino acid translation of SEQ ID NO: 1101 and 1098.

SEQ ID NO: 10 is sequence ID *rhg1\_peking\_amplicon* from the line *peking*, contains four *rhg1*, *v.1* exons at coding coordinates 113-264, 400-459, 1888-3710, and 3903-4501, and has an amino acid translation of SEQ ID NO: 1102.

SEQ ID NO: 11 is sequence ID *rhg1\_peking\_amplicon*, contains two *rhg1*, *v.2* exons at coding coordinates 1745-3710 and 3903-4501, and has an amino acid translation of SEQ ID NO: 1103.

SEQ ID NO: 12 is sequence ID *rhg1\_toyosuzu\_amplicon* from the line *toyosuzu*, contains four *rhg1*, *v.1* exons at coding coordinates 113-264, 400-459, 1890-3712, and 3924-4522, and has an amino acid translation of SEQ ID NO: 1104.

SEQ ID NO: 13 is sequence ID *rhg1\_toyosuzu\_amplicon*, contains two *rhg1*, *v.2* exons at coding coordinates 1747-3712 and 3924-4522, and has an amino acid translation of SEQ ID NO: 1105.

SEQ ID NO: 14 is sequence ID *rhg1\_will\_amplicon* from the line *will*, contains four *rhg1*, *v.1* exons at coding coordinates 113-264, 400-459, 1891-3713, and 3925-4523, and has an amino acid translation of SEQ ID NO: 1106.

SEQ ID NO: 15 is sequence ID *rhg1\_will\_amplicon*, contains two *rhg1*, *v.2* exons at coding coordinates 1748-3713 and 3925-4523, and has an amino acid translation of SEQ ID NO: 1107.

SEQ ID NO: 16 is sequence ID *rhg1\_a2704\_amplicon* from the line *A2704*, contains four *rhg1*, *v.1* exons at coding coordinates 113-264, 400-459, 1891-3713, and 3925-4523, and has an amino acid translation of SEQ ID NO: 1108.

SEQ ID NO: 17 is sequence ID *rhg1\_a2704\_amplicon*, contains two *rhg1*, *v.2* exons at coding coordinates 1748-3713 and 3925-4523, and has an amino acid translation of SEQ ID NO: 1109.

SEQ ID NO: 18 is sequence ID *rhg1\_noir\_amplicon* from the line *noir*, contains four *rhg1*, *v.1* exons at coding coordinates 113-264, 400-459, 1876-3698, and 3910-4508, and has an amino acid translation of SEQ ID NO: 1110.

SEQ ID NO: 19 is sequence ID *rhg1\_noir\_amplicon*, contains two *rhg1*, *v.2* exons at coding coordinates 1733-3698 and 3910-4508, and has an amino acid translation of SEQ ID NO: 1111.

SEQ ID NO: 20 is sequence ID *rhg1\_lee\_amplicon* from the line *lee*, contains four *rhg1*, *v.1* exons at coding coordinates 113-264, 400-459, 1876-3698, and 3910-4508, and has an amino acid translation of SEQ ID NO: 1112.

SEQ ID NO: 21 is sequence ID *rhg1\_lee\_amplicon*, contains two *rhg1*, *v.2* exons at coding coordinates 1733-3698 and 3910-4508, and has an amino acid translation of SEQ ID NO: 1113.

SEQ ID NO: 22 is sequence ID *rhg1\_pi200499\_amplicon* from the line *PI200499*, contains four *rhg1*, *v.1* exons at coding coordinates 113-264, 400-459, 1876-3698, and 3910-4508, and has an amino acid translation of SEQ ID NO: 1114.

SEQ ID NO: 23 is sequence ID rhg1\_pi200499\_amplicon, contains two *rhg1*, v.2 exons at coding coordinates 1733-3698 and 3910-4508, and has an amino acid translation of SEQ ID NO: 1115.

SEQ ID NO: 24 is sequence ID 240017\_region\_G3\_forward\_1, is a primer that hybridizes to coordinates 45051-45077 on contig 240017\_region\_G3 before the start codon, and can be used with SEQ ID NO: 25.

SEQ ID NO: 25 is sequence ID 240017\_region\_G3\_reverse\_1, is a primer that hybridizes to coordinates 47942-47918 on contig 240017\_region\_G3, and can be used with SEQ ID NO: 24.

SEQ ID NO: 26 is sequence ID 240017\_region\_G3\_forward\_2, is a primer that hybridizes to coordinates 47808-47831 on contig 240017\_region\_G3, and can be used with SEQ ID NO: 27.

SEQ ID NO: 27 is sequence ID 240017\_region\_G3\_reverse\_2, is a primer that hybridizes to coordinates 49553-49531 of contig 240017\_region\_G3 prior to the stop codon, and can be used with SEQ ID NO: 26.

Primers given by SEQ ID NOs: 24-27 are used to create the amplicons of SEQ ID NOs: 8-23. The final 22 bases are added to the actual amplicons in order to simulate the rest of the gene to the stop codon, in order to allow complete translation.

SEQ ID NO: 28 is sequence ID rhg1\_A3244\_amplicon\_cds, which is the coding sequence portion of SEQ ID NO: 8.

SEQ ID NO: 29 is sequence ID rhg1\_peking\_amplicon\_cds, which is the coding sequence portion of SEQ ID NO: 10.

SEQ ID NO: 30 is sequence ID rhg1\_toyosuzu\_amplicon\_cds, which is the coding sequence portion of SEQ ID NO: 12.

SEQ ID NO: 31 is sequence ID rhg1\_will\_amplicon\_cds, which is the coding sequence portion of SEQ ID NO: 14.

SEQ ID NO: 32 is sequence ID rhg1\_a2704\_amplicon\_cds, which is the coding sequence portion of SEQ ID NO: 16.

SEQ ID NO: 33 is sequence ID rhg1\_noir\_amplicon\_cds, which is the coding sequence portion of SEQ ID NO: 18.

SEQ ID NO: 34 is sequence ID rhg1\_lee\_amplicon\_cds, which is the coding sequence portion of SEQ ID NO: 20.

SEQ ID NO: 35 is sequence ID rhg1\_pi200499\_amplicon\_cds, which is the coding sequence portion of SEQ ID NO: 22.

SEQ ID NO: 36 is sequence ID rhg1\_A3244\_amplicon\_cds\_2, which is the coding sequence portion of SEQ ID NO: 9.

SEQ ID NO: 37 is sequence ID rhg1\_peking\_amplicon\_cds\_2, which is the coding sequence portion of SEQ ID NO: 11.

SEQ ID NO: 38 is sequence ID rhg1\_toyosuzu\_amplicon\_cds\_2, which is the coding sequence portion of SEQ ID NO: 13.

SEQ ID NO: 39 is sequence ID rhg1\_will\_amplicon\_cds\_2, which is the coding sequence portion of SEQ ID NO: 15.

SEQ ID NO: 40 is sequence ID rhg1\_a2704\_amplicon\_cds\_2, which is the coding sequence portion of SEQ ID NO: 17.

SEQ ID NO: 41 is sequence ID rhg1\_noir\_amplicon\_cds\_2, which is the coding sequence portion of SEQ ID NO: 19.

SEQ ID NO: 42 is sequence ID rhg1\_lee\_amplicon\_cds\_2, which is the coding sequence portion of SEQ ID NO: 21.

SEQ ID NO: 43 is sequence ID rhg1\_pi200499\_amplicon\_cds\_2, which is the coding sequence portion of SEQ ID NO: 23.

SEQ ID NOs: 44-53 and 1116-1119 all refer to *Rhg4* sequences

SEQ ID NO: 44 is sequence ID rhg4\_a3244\_amplicon from the line A3244, contains *Rhg4* at coding coordinates 79-2242 and 2958-3478, is made using SEQ ID NOs: 48 and 49, and has an amino acid translation of SEQ ID NO: 1116 and 1099.



SEQ ID NO: 45 is sequence ID rhg4\_Minsoy\_amplicon from the line Minsoy, contains *Rhg4* at coding coordinates 79-2242 and 2958-3478, is made using SEQ ID NOs: 48 and 49, and has an amino acid translation of SEQ ID NO: 1117.

SEQ ID NO: 46 is sequence ID rhg4\_Jack\_amplicon from the line Jack, contains *Rhg4* at coding coordinates 79-2242 and 2958-3478, is made using SEQ ID NO: 48 and 49, and has an amino acid translation of SEQ ID NO: 1118.

SEQ ID NO: 47 is sequence ID rhg4\_peking\_amplicon from the line Peking, contains *Rhg4* at coding coordinates 79-2242 and 2958-3478, is made using SEQ ID NOs: 48 and 49, and has an amino acid translation of SEQ ID NO: 1119.

SEQ ID NO: 48 is sequence ID 318O13\_region\_A3\_forward, hybridizes to coordinates 111727-111756 of contig 318O13\_region\_A3, and is a primer used with SEQ ID NO: 49 to create *Rhg4* amplicons.

SEQ ID NO: 49 is sequence ID 318O13\_region\_A3\_reverse, hybridizes to coordinates 115206-115177 of contig 318O13\_region\_A3, and is a primer used with SEQ ID NO: 48 to create *Rhg4* amplicons.

SEQ ID NO: 50 is sequence ID rhg4\_A3244\_amplicon\_cds, which is the coding sequence portion of SEQ ID NO: 44.

SEQ ID NO: 51 is sequence ID rhg4\_Minsoy\_amplicon\_cds, which is the coding sequence portion of SEQ ID NO: 45.

SEQ ID NO: 52 is sequence ID rhg4\_Jack\_amplicon\_cds, which is the coding sequence portion of SEQ ID NO: 46.

SEQ ID NO: 53 is sequence ID rhg4\_peking\_amplicon\_cds, which is the coding sequence portion of SEQ ID NO: 47.

SEQ ID NO: 1120 is sequence ID consensusLRR, which is a consensus sequence for the LRR repeats shown in Figures 1 and 2.

SEQ ID NO: 1121 is sequence ID rhg1LRR, which is the amino acid sequence of the LRR domain shown in Figure 1.

SEQ ID NO: 1122 is sequence ID Rhg4LRR, which is the amino acid sequence of the LRR domain shown in Figure 2.

SEQ ID NO: 1123 is sequence ID 240017\_region\_G3\_forward\_1\_b, which is an alternate primer that hybridizes to coordinates 45046-45072 on config 240017\_region\_G3 before the start codon, and which can be used with SEQ ID NO: 25.

Table 1 below provides further information on the sequences described herein.

In table 1, for all rows, "Seq Num" refers to the corresponding SEQ ID NO in the sequence listing.

For rows with SEQ ID NOs: 1-53 and 1120-1123 "Seq ID" refers to the name of the SEQ ID NO given in the "Seq Num" column.

For rows with SEQ ID NOs: 2-4, 8-23, and 44-47 "Coding Sequence" refers to the coordinates of the coding portion of the SEQ ID NO given in the "Seq Num" column, and "AA" refers to the SEQ ID NO that is the amino acid translation of the SEQ ID NO given in the "Seq Num" column.

For rows with SEQ ID NOs: 24-27 and 1123, "Primer location on 240017\_region\_G3" refers to the coordinates of the 240017\_region\_G3 config to which the SEQ ID NO given in the "Seq Num" column hybridizes.

For rows with SEQ ID NOs: 48 and 49, "Primer location on 318O13\_region\_A3" refers to the coordinates of the 318O13\_region\_A3 config to which the SEQ ID NO given in the "Seq Num" column hybridizes.

For rows with SEQ ID NOs: 54-400, "Seq ID" refers to the names of amplicon sequences. Within the Seq ID is the "\_\_\_" (double length underscore) symbol. The name before this symbol refers to the name of the config in which the amplicon is found, and the numbers after this symbol refer to the nucleotide location of the SSR on the config.

For rows with SEQ ID NOs: 401-1096, "Seq ID" refers to the names of primer sequences used in PCR to generate the amplicon sequences in table 1. For these rows, the "Seq ID" name contains the same name as the amplicon that is generated by the pair of primers of which the SEQ ID NO referred to in the first column is a member. The "Seq

ID" name also contains either "Forward" or "Reverse," which indicates the orientation of the primer. For these sequences, "location of primer on contig start" and "location of primer on contig end" refer, respectively, to the first and last base number of the contig on which the primer aligns.

TABLE 1

Seq Num	Seq ID		
1	515O02_region_G2		

Seq Num	Seq ID	Coding Sequence	AA No.
2	240O17_region_G3	45163-45314,45450-45509,46941-48763,48975-49573	1097
3	240O17_region_G3	46798-48763,48975-49573	1098
4	318O13_region_A3	111805-113968,114684-115204	1099

Seq Num	Seq ID		
5	240O17_region_G3_8_mRNA		
6	240O17_region_G3_8_cds		
7	318O13_region_A3_17_cds		

Seq Num	Seq ID	Coding Sequence	AA No.
8	rhg1_A3244_amplicon	113-264,400-459,1891-3713,3925-4523	1100
9	rhg1_A3244_amplicon	1748-3713,3925-4523	1101
10	rhg1_peking_amplicon	113-264,400-459,1888-3710,3903-4501	1102
11	rhg1_peking_amplicon	1745-3710,3903-4501	1103
12	rhg1_toyosuzu_amplicon	113-264,400-459,1890-3712,3924-4522	1104
13	rhg1_toyosuzu_amplicon	1747-3712,3924-4522	1105
14	rhg1_will_amplicon	113-264,400-459,1891-3713,3925-4523	1106
15	rhg1_will_amplicon	1748-3713,3925-4523	1107
16	rhg1_a2704_amplicon	113-264,400-459,1891-3713,3925-4523	1108
17	rhg1_a2704_amplicon	1748-3713,3925-4523	1109
18	rhg1_noir_amplicon	113-264,400-459,1876-3698,3910-4508	1110
19	rhg1_noir_amplicon	1733-3698,3910-4508	1111
20	rhg1_lee_amplicon	113-264,400-459,1876-3698,3910-4508	1112
21	rhg1_lee_amplicon	1733-3698,3910-4508	1113
22	rhg1_pi200499_amplicon	113-264,400-459,1876-3698,3910-4508	1114
23	rhg1_pi200499_amplicon	1733-3698,3910-4508	1115

Seq Num	Seq ID	Primer location on 240O17_region_G3	
24	240O17_region_G3_forward_1	45051-45077	
25	240O17_region_G3_reverse_1	47942-47918	
26	240O17_region_G3_forward_2	47808-47831	
27	240O17_region_G3_reverse_2	49553-49531	

Seq Num	Seq ID		
28	rhg1_A3244_amplicon_cds		
29	rhg1_peking_amplicon_cds		
30	rhg1_toyosuzu_amplicon_cds		
31	rhg1_will_amplicon_cds		

Seq Num	Seq ID		
32	rhg1_a2704_amplicon_cds		
33	rhg1_noir_amplicon_cds		
34	rhg1_lee_amplicon_cds		
35	rhg1_pi200499_amplicon_cds		
36	rhg1_A3244_amplicon_cds_2		
37	rhg1_peking_amplicon_cds_2		
38	rhg1_toyosuzu_amplicon_cds_2		
39	rhg1_will_amplicon_cds_2		
40	rhg1_a2704_amplicon_cds_2		
41	rhg1_noir_amplicon_cds_2		
42	rhg1_lee_amplicon_cds_2		
43	rhg1_pi200499_amplicon_cds_2		

Seq Num	Seq ID	Coding Sequence	AA No.
44	rhg4_a3244_amplicon	79-2242,2958-3478	1116
45	rhg4_Minsoy_amplicon	79-2242,2958-3478	1117
46	rhg4_Jack_amplicon	79-2242,2958-3478	1118
47	rhg4_peking_amplicon	79-2242,2958-3478	1119

Seq Num	Seq ID	Primer location on 318O13_region_A3	
48	318O13_region_A3_forward	111727-111756	
49	318O13_region_A3_reverse	115206-115177	

Seq Num	Seq ID		
50	rhg4_A3244_amplicon_cds		
51	rhg4_Minsoy_amplicon_cds		
52	rhg4_Jack_amplicon_cds		
53	rhg4_peking_amplicon_cds		

Seq Num	Seq ID		
54	240O17_region_G3_289711_11		
55	240O17_region_G3_236585_14		
56	240O17_region_G3_168772_13		
57	240O17_region_G3_332420_21		
58	240O17_region_G3_228126_18		
59	240O17_region_G3_139723_11		
60	240O17_region_G3_280585_14		
61	240O17_region_G3_70509_14		
62	240O17_region_G3_50537_17		
63	240O17_region_G3_231556_17		
64	240O17_region_G3_117057_11		
65	240O17_region_G3_23092_13		
66	240O17_region_G3_297741_14		
67	240O17_region_G3_206502_14		
68	240O17_region_G3_221223_13		
69	240O17_region_G3_169084_14		
70	240O17_region_G3_94891_14		
71	240O17_region_G3_281852_61		
72	240O17_region_G3_46583_12		
73	240O17_region_G3_306835_13		
74	240O17_region_G3_85471_12		
75	240O17_region_G3_257208_12		

Seq Num	Seq ID		
76	240O17_region_G3_150390_17		
77	240O17_region_G3_34697_75		
78	240O17_region_G3_150374_13		
79	240O17_region_G3_40513_22		
80	240O17_region_G3_268602_14		
81	240O17_region_G3_25357_13		
82	240O17_region_G3_137548_13		
83	240O17_region_G3_139131_13		
84	240O17_region_G3_203855_12		
85	240O17_region_G3_199049_15		
86	240O17_region_G3_320907_12		
87	240O17_region_G3_16407_17		
88	240O17_region_G3_206516_17		
89	240O17_region_G3_264495_13		
90	240O17_region_G3_156785_13		
91	240O17_region_G3_187129_12		
92	240O17_region_G3_214106_13		
93	240O17_region_G3_149013_12		
94	240O17_region_G3_326352_16		
95	240O17_region_G3_278962_12		
96	240O17_region_G3_256930_13		
97	240O17_region_G3_29646_14		
98	240O17_region_G3_29618_13		
99	240O17_region_G3_108561_14		
100	240O17_region_G3_143975_14		
101	240O17_region_G3_108431_20		
102	240O17_region_G3_281764_11		
103	240O17_region_G3_130058_15		
104	240O17_region_G3_310590_52		
105	240O17_region_G3_313405_14		
106	240O17_region_G3_302190_13		
107	240O17_region_G3_225343_17		
108	240O17_region_G3_208823_14		
109	240O17_region_G3_74285_11		
110	240O17_region_G3_109052_16		
111	240O17_region_G3_6395_12		
112	240O17_region_G3_244905_16		
113	240O17_region_G3_244956_13		
114	240O17_region_G3_117220_13		
115	240O17_region_G3_134707_14		
116	240O17_region_G3_35078_13		
117	240O17_region_G3_210506_16		
118	240O17_region_G3_116961_26		
119	240O17_region_G3_51073_13		
120	240O17_region_G3_55291_15		
121	240O17_region_G3_229651_18		
122	240O17_region_G3_303308_19		
123	240O17_region_G3_168373_20		
124	240O17_region_G3_253333_17		
125	240O17_region_G3_5791_13		
126	240O17_region_G3_206841_19		
127	240O17_region_G3_202827_12		
128	240O17_region_G3_322656_13		

Seq Num	Seq ID		
129	240017_region_G3_111841_14		
130	240017_region_G3_192719_13		
131	240017_region_G3_195630_17		
132	240017_region_G3_69999_13		
133	240017_region_G3_11176_13		
134	240017_region_G3_228643_13		
135	240017_region_G3_88478_19		
136	240017_region_G3_108950_13		
137	240017_region_G3_121054_14		
138	240017_region_G3_188337_14		
139	240017_region_G3_255944_21		
140	240017_region_G3_219518_14		
141	240017_region_G3_235601_15		
142	240017_region_G3_301529_13		
143	240017_region_G3_94795_14		
144	240017_region_G3_46703_23		
145	240017_region_G3_59616_14		
146	240017_region_G3_296933_15		
147	240017_region_G3_192428_17		
148	240017_region_G3_191490_14		
149	240017_region_G3_201115_11		
150	240017_region_G3_72882_15		
151	240017_region_G3_69514_13		
152	240017_region_G3_37699_47		
153	240017_region_G3_11301_29		
154	240017_region_G3_141875_12		
155	240017_region_G3_98090_18		
156	240017_region_G3_43298_35		
157	240017_region_G3_262094_11		
158	240017_region_G3_262079_15		
159	240017_region_G3_59090_12		
160	240017_region_G3_245723_13		
161	240017_region_G3_194628_54		
162	240017_region_G3_4566_16		
163	240017_region_G3_96209_14		
164	240017_region_G3_248715_17		
165	240017_region_G3_71410_40		
166	240017_region_G3_226519_13		
167	240017_region_G3_11282_19		
168	240017_region_G3_170504_12		
169	240017_region_G3_40864_14		
170	240017_region_G3_13529_14		
171	240017_region_G3_22858_14		
172	240017_region_G3_309211_13		
173	240017_region_G3_55568_26		
174	240017_region_G3_73238_16		
175	240017_region_G3_52488_19		
176	318O13_region_A3_471518_14		
177	318O13_region_A3_231599_23		
178	318O13_region_A3_375912_13		
179	318O13_region_A3_180013_12		
180	318O13_region_A3_171606_14		
181	318O13_region_A3_416256_13		

Seq Num	Seq ID		
182	318O13_region_A3_231395_15		
183	318O13_region_A3_5502_47		
184	318O13_region_A3_93061_14		
185	318O13_region_A3_111684_19		
186	318O13_region_A3_69328_14		
187	318O13_region_A3_36529_17		
188	318O13_region_A3_139128_12		
189	318O13_region_A3_495674_13		
190	318O13_region_A3_187577_13		
191	318O13_region_A3_453036_14		
192	318O13_region_A3_374041_13		
193	318O13_region_A3_3412_11		
194	318O13_region_A3_276495_28		
195	318O13_region_A3_151839_17		
196	318O13_region_A3_292912_12		
197	318O13_region_A3_104560_12		
198	318O13_region_A3_65193_11		
199	318O13_region_A3_110573_70		
200	318O13_region_A3_65117_12		
201	318O13_region_A3_490837_16		
202	318O13_region_A3_107448_11		
203	318O13_region_A3_331_23		
204	318O13_region_A3_193470_13		
205	318O13_region_A3_183305_14		
206	318O13_region_A3_55050_14		
207	318O13_region_A3_224693_21		
208	318O13_region_A3_207216_12		
209	318O13_region_A3_4654_22		
210	318O13_region_A3_408959_13		
211	318O13_region_A3_132288_22		
212	318O13_region_A3_292822_20		
213	318O13_region_A3_311076_12		
214	318O13_region_A3_509623_13		
215	318O13_region_A3_190404_14		
216	318O13_region_A3_164916_15		
217	318O13_region_A3_21028_13		
218	318O13_region_A3_208012_17		
219	318O13_region_A3_484089_14		
220	318O13_region_A3_332780_17		
221	318O13_region_A3_480137_37		
222	318O13_region_A3_441056_14		
223	318O13_region_A3_77486_11		
224	318O13_region_A3_272468_11		
225	318O13_region_A3_425319_17		
226	318O13_region_A3_413879_31		
227	318O13_region_A3_80477_64		
228	318O13_region_A3_277272_50		
229	318O13_region_A3_509642_13		
230	318O13_region_A3_321771_14		
231	318O13_region_A3_26788_12		
232	318O13_region_A3_262706_16		
233	318O13_region_A3_243928_16		
234	318O13_region_A3_23246_14		

Seq Num	Seq ID		
235	318O13_region_A3_165406_12		
236	318O13_region_A3_486294_14		
237	318O13_region_A3_46754_12		
238	318O13_region_A3_381116_15		
239	318O13_region_A3_350369_11		
240	318O13_region_A3_138841_13		
241	318O13_region_A3_12158_14		
242	318O13_region_A3_315368_13		
243	318O13_region_A3_307549_13		
244	318O13_region_A3_159857_14		
245	318O13_region_A3_140551_15		
246	318O13_region_A3_279869_11		
247	318O13_region_A3_78292_35		
248	318O13_region_A3_185019_12		
249	318O13_region_A3_409164_13		
250	318O13_region_A3_75392_14		
251	318O13_region_A3_231320_12		
252	318O13_region_A3_381102_14		
253	318O13_region_A3_491826_15		
254	318O13_region_A3_56365_21		
255	318O13_region_A3_372628_15		
256	318O13_region_A3_302609_11		
257	318O13_region_A3_341804_11		
258	318O13_region_A3_217037_11		
259	318O13_region_A3_264929_68		
260	318O13_region_A3_55499_12		
261	318O13_region_A3_295634_14		
262	318O13_region_A3_269358_15		
263	318O13_region_A3_457009_24		
264	318O13_region_A3_176598_14		
265	318O13_region_A3_278266_12		
266	318O13_region_A3_391810_12		
267	318O13_region_A3_269485_15		
268	318O13_region_A3_359247_17		
269	318O13_region_A3_315094_13		
270	318O13_region_A3_307823_13		
271	318O13_region_A3_248588_15		
272	318O13_region_A3_252426_85		
273	318O13_region_A3_513314_16		
274	318O13_region_A3_68183_14		
275	318O13_region_A3_471191_13		
276	318O13_region_A3_163547_18		
277	318O13_region_A3_417867_15		
278	318O13_region_A3_332465_14		
279	318O13_region_A3_207697_14		
280	318O13_region_A3_277229_43		
281	318O13_region_A3_36366_11		
282	318O13_region_A3_91970_12		
283	318O13_region_A3_211533_11		
284	318O13_region_A3_336301_11		
285	318O13_region_A3_441603_14		
286	318O13_region_A3_468354_15		
287	318O13_region_A3_188983_18		



Seq Num	Seq ID		
288	318O13_region_A3_115502_17		
289	318O13_region_A3_163006_13		
290	318O13_region_A3_119283_14		
291	318O13_region_A3_491126_11		
292	318O13_region_A3_99512_21		
293	318O13_region_A3_280291_17		
294	318O13_region_A3_138443_19		
295	318O13_region_A3_115973_14		
296	318O13_region_A3_329977_14		
297	318O13_region_A3_205203_14		
298	318O13_region_A3_153114_12		
299	318O13_region_A3_34581_13		
300	318O13_region_A3_292577_19		
301	318O13_region_A3_445391_20		
302	318O13_region_A3_350540_17		
303	318O13_region_A3_453879_15		
304	318O13_region_A3_201246_13		
305	318O13_region_A3_326020_13		
306	318O13_region_A3_503801_14		
307	318O13_region_A3_302400_52		
308	318O13_region_A3_448857_15		
309	318O13_region_A3_48364_14		
310	318O13_region_A3_251804_48		
311	318O13_region_A3_382583_13		
312	318O13_region_A3_124737_14		
313	318O13_region_A3_124766_13		
314	318O13_region_A3_461351_16		
315	318O13_region_A3_64953_19		
316	318O13_region_A3_366586_13		
317	318O13_region_A3_46190_15		
318	318O13_region_A3_81016_11		
319	318O13_region_A3_134426_14		
320	318O13_region_A3_292724_14		
321	318O13_region_A3_187096_17		
322	318O13_region_A3_381693_13		
323	318O13_region_A3_361286_33		
324	318O13_region_A3_482668_14		
325	318O13_region_A3_128002_12		
326	318O13_region_A3_499270_14		
327	318O13_region_A3_231650_12		
328	318O13_region_A3_199851_13		
329	318O13_region_A3_324629_13		
330	318O13_region_A3_374190_19		
331	318O13_region_A3_460603_13		
332	318O13_region_A3_108681_14		
333	318O13_region_A3_459791_47		
334	318O13_region_A3_4257_20		
335	318O13_region_A3_238810_14		
336	318O13_region_A3_245817_14		
337	318O13_region_A3_245956_14		
338	318O13_region_A3_74148_14		
339	318O13_region_A3_74089_15		
340	318O13_region_A3_241686_12		

Seq Num	Seq ID		
341	318O13_region_A3_47476_12		
342	318O13_region_A3_164550_12		
343	318O13_region_A3_101255_15		
344	515O02_region_G2_16189_11		
345	515O02_region_G2_71925_13		
346	515O02_region_G2_4707_12		
347	515O02_region_G2_118904_18		
348	515O02_region_G2_13655_17		
349	515O02_region_G2_53900_13		
350	515O02_region_G2_8079_14		
351	515O02_region_G2_9969_28		
352	515O02_region_G2_72308_77		
353	515O02_region_G2_99475_19		
354	515O02_region_G2_118615_18		
355	515O02_region_G2_119001_46		
356	515O02_region_G2_118958_43		
357	515O02_region_G2_17197_13		
358	515O02_region_G2_105163_29		
359	515O02_region_G2_111335_13		
360	515O02_region_G2_106396_13		
361	515O02_region_G2_59229_17		
362	515O02_region_G2_73795_20		
363	515O02_region_G2_85664_20		
364	515O02_region_G2_36921_17		
365	515O02_region_G2_124150_19		
366	515O02_region_G2_5089_14		
367	515O02_region_G2_58221_15		
368	515O02_region_G2_96139_14		
369	515O02_region_G2_70595_13		
370	515O02_region_G2_4340_15		
371	515O02_region_G2_90417_11		
372	515O02_region_G2_49711_17		
373	515O02_region_G2_63053_13		
374	515O02_region_G2_63076_14		
375	515O02_region_G2_44442_12		
376	515O02_region_G2_44422_19		
377	515O02_region_G2_44158_19		
378	515O02_region_G2_44141_17		
379	515O02_region_G2_90762_17		
380	515O02_region_G2_106241_14		
381	515O02_region_G2_109676_12		
382	515O02_region_G2_86242_14		
383	515O02_region_G2_83109_12		
384	515O02_region_G2_10461_15		
385	515O02_region_G2_67608_15		
386	515O02_region_G2_63275_46		
387	515O02_region_G2_62405_14		
388	515O02_region_G2_33563_12		
389	515O02_region_G2_33146_14		
390	515O02_region_G2_102179_29		
391	515O02_region_G2_2646_15		
392	515O02_region_G2_76652_24		
393	515O02_region_G2_66280_14		

Seq Num	Seq ID		
394	515002_region_G2_54768_13		
395	515002_region_G2_62580_14		
396	515002_region_G2_34598_55		
397	515002_region_G2_77680_13		
398	515002_region_G2_77693_12		
399	515002_region_G2_97392_14		
400	515002_region_G2_97359_15		

Seq Num	Seq ID	location of primer on config start	location of primer on config end
401	240017_region_G3_289711_11_Forward_Primer	289637	289661
402	240017_region_G3_289711_11_Reverse_Primer	289756	289732
403	240017_region_G3_236585_14_Forward_Primer	236511	236535
404	240017_region_G3_236585_14_Reverse_Primer	236638	236614
405	240017_region_G3_168772_13_Forward_Primer	168683	168707
406	240017_region_G3_168772_13_Reverse_Primer	168811	168786
407	240017_region_G3_332420_21_Forward_Primer	332375	332399
408	240017_region_G3_332420_21_Reverse_Primer	332505	332481
409	240017_region_G3_228126_18_Forward_Primer	228048	228072
410	240017_region_G3_228126_18_Reverse_Primer	228182	228158
411	240017_region_G3_139723_11_Forward_Primer	139666	139690
412	240017_region_G3_139723_11_Reverse_Primer	139802	139778
413	240017_region_G3_280585_14_Forward_Primer	280524	280550
414	240017_region_G3_280585_14_Reverse_Primer	280661	280637
415	240017_region_G3_70509_14_Forward_Primer	70478	70502
416	240017_region_G3_70509_14_Reverse_Primer	70616	70592
417	240017_region_G3_50537_17_Forward_Primer	50455	50479
418	240017_region_G3_50537_17_Reverse_Primer	50593	50569
419	240017_region_G3_231556_17_Forward_Primer	231468	231492
420	240017_region_G3_231556_17_Reverse_Primer	231606	231582
421	240017_region_G3_117057_11_Forward_Primer	117029	117053
422	240017_region_G3_117057_11_Reverse_Primer	117169	117145
423	240017_region_G3_23092_13_Forward_Primer	23010	23034
424	240017_region_G3_23092_13_Reverse_Primer	23151	23127
425	240017_region_G3_297741_14_Forward_Primer	297680	297704
426	240017_region_G3_297741_14_Reverse_Primer	297823	297799
427	240017_region_G3_206502_14_Forward_Primer	206456	206480
428	240017_region_G3_206502_14_Reverse_Primer	206600	206581
429	240017_region_G3_221223_13_Forward_Primer	221134	221158
430	240017_region_G3_221223_13_Reverse_Primer	221278	221254
431	240017_region_G3_169084_14_Forward_Primer	169051	169075
432	240017_region_G3_169084_14_Reverse_Primer	169196	169173
433	240017_region_G3_94891_14_Forward_Primer	94784	94808
434	240017_region_G3_94891_14_Reverse_Primer	94929	94905
435	240017_region_G3_7439_12_Forward_Primer	7397	7421
436	240017_region_G3_7439_12_Reverse_Primer	7542	7518
437	240017_region_G3_281852_61_Forward_Primer	281797	281821
438	240017_region_G3_281852_61_Reverse_Primer	281943	281919
439	240017_region_G3_46583_12_Forward_Primer	46554	46578
440	240017_region_G3_46583_12_Reverse_Primer	46700	46676
441	240017_region_G3_306835_13_Forward_Primer	306727	306751
442	240017_region_G3_306835_13_Reverse_Primer	306874	306849

Seq Num	Seq ID	location of primer on contig start	location of primer on contig end
443	240017_region_G3_85471_12_Forward_Primer	85359	85383
444	240017_region_G3_85471_12_Reverse_Primer	85507	85483
445	240017_region_G3_257208_12_Forward_Primer	257129	257153
446	240017_region_G3_257208_12_Reverse_Primer	257278	257254
447	240017_region_G3_150390_17_Forward_Primer	150327	150351
448	240017_region_G3_150390_17_Reverse_Primer	150476	150452
449	240017_region_G3_34697_75_Forward_Primer	34662	34685
450	240017_region_G3_34697_75_Reverse_Primer	34811	34787
451	240017_region_G3_150374_13_Forward_Primer	150327	150351
452	240017_region_G3_150374_13_Reverse_Primer	150476	150452
453	240017_region_G3_40513_22_Forward_Primer	40422	40446
454	240017_region_G3_40513_22_Reverse_Primer	40572	40548
455	240017_region_G3_268602_14_Forward_Primer	268555	268579
456	240017_region_G3_268602_14_Reverse_Primer	268705	268681
457	240017_region_G3_25357_13_Forward_Primer	25271	25295
458	240017_region_G3_25357_13_Reverse_Primer	25422	25402
459	240017_region_G3_137548_13_Forward_Primer	139088	139111
459	240017_region_G3_137548_13_Forward_Primer	137505	137528
460	240017_region_G3_137548_13_Reverse_Primer	139239	139215
460	240017_region_G3_137548_13_Reverse_Primer	137656	137632
461	240017_region_G3_139131_13_Forward_Primer	139088	139111
462	240017_region_G3_139131_13_Reverse_Primer	139239	139215
463	240017_region_G3_203855_12_Forward_Primer	203749	203773
464	240017_region_G3_203855_12_Reverse_Primer	203901	203877
465	240017_region_G3_199049_15_Forward_Primer	199008	199033
466	240017_region_G3_199049_15_Reverse_Primer	199160	199136
467	240017_region_G3_320907_12_Forward_Primer	320885	320906
468	240017_region_G3_320907_12_Reverse_Primer	321038	321015
469	240017_region_G3_16407_17_Forward_Primer	16330	16354
470	240017_region_G3_16407_17_Reverse_Primer	16483	16459
471	240017_region_G3_206516_17_Forward_Primer	206482	206506
472	240017_region_G3_206516_17_Reverse_Primer	206635	206616
473	240017_region_G3_264495_13_Forward_Primer	264423	264447
474	240017_region_G3_264495_13_Reverse_Primer	264577	264553
475	240017_region_G3_156785_13_Forward_Primer	156713	156737
476	240017_region_G3_156785_13_Reverse_Primer	156868	156844
477	240017_region_G3_187129_12_Forward_Primer	187068	187092
478	240017_region_G3_187129_12_Reverse_Primer	187223	187199
479	240017_region_G3_214106_13_Forward_Primer	214042	214067
480	240017_region_G3_214106_13_Reverse_Primer	214197	214173
481	240017_region_G3_149013_12_Forward_Primer	148898	148922
482	240017_region_G3_149013_12_Reverse_Primer	149053	149027
483	240017_region_G3_326352_16_Forward_Primer	326311	326335
484	240017_region_G3_326352_16_Reverse_Primer	326467	326443
485	240017_region_G3_278962_12_Forward_Primer	278933	278957
486	240017_region_G3_278962_12_Reverse_Primer	279089	279065
487	240017_region_G3_256930_13_Forward_Primer	256850	256874
488	240017_region_G3_256930_13_Reverse_Primer	257006	256982
489	240017_region_G3_29646_14_Forward_Primer	29589	29613
490	240017_region_G3_29646_14_Reverse_Primer	29746	29721
491	240017_region_G3_29618_13_Forward_Primer	29589	29613

Seq Num	Seq ID	location of primer on config start	location of primer on config end
492	240017_region_G3_29618_13 Reverse_Primer	29746	29721
493	240017_region_G3_108561_14 Forward_Primer	108518	108542
494	240017_region_G3_108561_14 Reverse_Primer	108675	108651
495	240017_region_G3_143975_14 Forward_Primer	143939	143964
496	240017_region_G3_143975_14 Reverse_Primer	144096	144072
497	240017_region_G3_108431_20 Forward_Primer	108362	108386
498	240017_region_G3_108431_20 Reverse_Primer	108520	108497
499	240017_region_G3_281764_11 Forward_Primer	281645	281669
500	240017_region_G3_281764_11 Reverse_Primer	281803	281779
501	240017_region_G3_130058_15 Forward_Primer	129994	130018
502	240017_region_G3_130058_15 Reverse_Primer	130153	130129
503	240017_region_G3_310590_52 Forward_Primer	310533	310557
504	240017_region_G3_310590_52 Reverse_Primer	310692	310668
505	240017_region_G3_313405_14 Forward_Primer	313345	313369
506	240017_region_G3_313405_14 Reverse_Primer	313505	313481
507	240017_region_G3_302190_13 Forward_Primer	302093	302119
508	240017_region_G3_302190_13 Reverse_Primer	302253	302229
509	240017_region_G3_225343_17 Forward_Primer	225315	225338
510	240017_region_G3_225343_17 Reverse_Primer	225475	225451
511	240017_region_G3_208823_14 Forward_Primer	208760	208784
512	240017_region_G3_208823_14 Reverse_Primer	208921	208897
513	240017_region_G3_74285_11 Forward_Primer	74220	74244
514	240017_region_G3_74285_11 Reverse_Primer	74382	74358
515	240017_region_G3_109052_16 Forward_Primer	108999	109023
516	240017_region_G3_109052_16 Reverse_Primer	109161	109137
517	240017_region_G3_6395_12 Forward_Primer	6285	6309
518	240017_region_G3_6395_12 Reverse_Primer	6447	6423
519	240017_region_G3_244905_16 Forward_Primer	244865	244890
520	240017_region_G3_244905_16 Reverse_Primer	245028	245004
521	240017_region_G3_244956_13 Forward_Primer	244865	244890
522	240017_region_G3_244956_13 Reverse_Primer	245028	245004
523	240017_region_G3_117220_13 Forward_Primer	117175	117199
524	240017_region_G3_117220_13 Reverse_Primer	117339	117315
525	240017_region_G3_134707_14 Forward_Primer	134584	134608
526	240017_region_G3_134707_14 Reverse_Primer	134749	134725
527	240017_region_G3_35078_13 Forward_Primer	34990	35013
528	240017_region_G3_35078_13 Reverse_Primer	35157	35133
529	240017_region_G3_210506_16 Forward_Primer	210477	210501
530	240017_region_G3_210506_16 Reverse_Primer	210644	210620
531	240017_region_G3_116961_26 Forward_Primer	116885	116909
532	240017_region_G3_116961_26 Reverse_Primer	117053	117029
533	240017_region_G3_51073_13 Forward_Primer	50979	51003
534	240017_region_G3_51073_13 Reverse_Primer	51147	51123
535	240017_region_G3_55291_15 Forward_Primer	55164	55188
536	240017_region_G3_55291_15 Reverse_Primer	55333	55309
537	240017_region_G3_229651_18 Forward_Primer	229615	229639
538	240017_region_G3_229651_18 Reverse_Primer	229784	229760
539	240017_region_G3_303308_19 Forward_Primer	303284	303307
540	240017_region_G3_303308_19 Reverse_Primer	303454	303429
541	240017_region_G3_168373_20 Forward_Primer	168262	168286
542	240017_region_G3_168373_20 Reverse_Primer	168432	168408

Seq Num	Seq ID	location of primer on config start	location of primer on config end
543	240017_region_G3_253333_17_Forward_Primer	253257	253281
544	240017_region_G3_253333_17_Reverse_Primer	253428	253404
545	240017_region_G3_5791_13_Forward_Primer	5766	5790
546	240017_region_G3_5791_13_Reverse_Primer	5937	5912
547	240017_region_G3_206841_19_Forward_Primer	206821	206840
548	240017_region_G3_206841_19_Reverse_Primer	206993	206969
549	240017_region_G3_202827_12_Forward_Primer	202782	202806
550	240017_region_G3_202827_12_Reverse_Primer	202956	202932
551	240017_region_G3_322656_13_Forward_Primer	322572	322598
552	240017_region_G3_322656_13_Reverse_Primer	322748	322724
553	240017_region_G3_111841_14_Forward_Primer	111709	111733
554	240017_region_G3_111841_14_Reverse_Primer	111886	111861
555	240017_region_G3_192719_13_Forward_Primer	192589	192613
556	240017_region_G3_192719_13_Reverse_Primer	192767	192743
557	240017_region_G3_195630_17_Forward_Primer	195490	195514
558	240017_region_G3_195630_17_Reverse_Primer	195672	195648
559	240017_region_G3_69999_13_Forward_Primer	69858	69881
560	240017_region_G3_69999_13_Reverse_Primer	70040	70016
561	240017_region_G3_11176_13_Forward_Primer	11060	11084
562	240017_region_G3_11176_13_Reverse_Primer	11243	11219
563	240017_region_G3_228643_13_Forward_Primer	228529	228553
564	240017_region_G3_228643_13_Reverse_Primer	228713	228689
565	240017_region_G3_88478_19_Forward_Primer	88378	88402
566	240017_region_G3_88478_19_Reverse_Primer	88562	88538
567	240017_region_G3_108950_13_Forward_Primer	108838	108858
568	240017_region_G3_108950_13_Reverse_Primer	109023	108998
569	240017_region_G3_121054_14_Forward_Primer	120911	120935
570	240017_region_G3_121054_14_Reverse_Primer	121096	121072
571	240017_region_G3_188337_14_Forward_Primer	188204	188228
572	240017_region_G3_188337_14_Reverse_Primer	191544	191520
572	240017_region_G3_188337_14_Reverse_Primer	188391	188367
573	240017_region_G3_255944_21_Forward_Primer	255879	255903
574	240017_region_G3_255944_21_Reverse_Primer	256068	256044
575	240017_region_G3_219518_14_Forward_Primer	219420	219444
576	240017_region_G3_219518_14_Reverse_Primer	219609	219585
577	240017_region_G3_235601_15_Forward_Primer	235483	235507
578	240017_region_G3_235601_15_Reverse_Primer	235673	235649
579	240017_region_G3_301529_13_Forward_Primer	301498	301522
580	240017_region_G3_301529_13_Reverse_Primer	301689	301665
581	240017_region_G3_94795_14_Forward_Primer	94735	94756
582	240017_region_G3_94795_14_Reverse_Primer	94929	94905
583	240017_region_G3_46703_23_Forward_Primer	46676	46700
584	240017_region_G3_46703_23_Reverse_Primer	46870	46846
585	240017_region_G3_59616_14_Forward_Primer	59539	59563
586	240017_region_G3_59616_14_Reverse_Primer	59738	59714
587	240017_region_G3_296933_15_Forward_Primer	296908	296932
588	240017_region_G3_296933_15_Reverse_Primer	297113	297089
589	240017_region_G3_192428_17_Forward_Primer	192402	192426
590	240017_region_G3_192428_17_Reverse_Primer	192613	192589
591	240017_region_G3_191490_14_Forward_Primer	191332	191356
592	240017_region_G3_191490_14_Reverse_Primer	191544	191520

Seq Num	Seq ID	location of primer on config start	location of primer on config end
593	240O17_region_G3_201115_11_Forward_Primer	200994	201018
594	240O17_region_G3_201115_11_Reverse_Primer	201214	201189
595	240O17_region_G3_72882_15_Forward_Primer	72848	72874
596	240O17_region_G3_72882_15_Reverse_Primer	73068	73042
597	240O17_region_G3_69514_13_Forward_Primer	69411	69437
598	240O17_region_G3_69514_13_Reverse_Primer	69632	69608
599	240O17_region_G3_37699_47_Forward_Primer	37601	37625
600	240O17_region_G3_37699_47_Reverse_Primer	37827	37802
601	240O17_region_G3_11301_29_Forward_Primer	11274	11300
602	240O17_region_G3_11301_29_Reverse_Primer	11501	11477
603	240O17_region_G3_141875_12_Forward_Primer	141729	141750
604	240O17_region_G3_141875_12_Reverse_Primer	141964	141939
605	240O17_region_G3_98090_18_Forward_Primer	98037	98062
606	240O17_region_G3_98090_18_Reverse_Primer	98274	98250
607	240O17_region_G3_43298_35_Forward_Primer	43144	43168
608	240O17_region_G3_43298_35_Reverse_Primer	43387	43363
609	240O17_region_G3_262094_11_Forward_Primer	261989	262014
610	240O17_region_G3_262094_11_Reverse_Primer	262236	262211
611	240O17_region_G3_262079_15_Forward_Primer	261989	262014
612	240O17_region_G3_262079_15_Reverse_Primer	262236	262211
613	240O17_region_G3_59090_12_Forward_Primer	58986	59012
614	240O17_region_G3_59090_12_Reverse_Primer	59248	59224
615	240O17_region_G3_245723_13_Forward_Primer	245502	245526
616	240O17_region_G3_245723_13_Reverse_Primer	245766	245742
617	240O17_region_G3_194628_54_Forward_Primer	194581	194607
618	240O17_region_G3_194628_54_Reverse_Primer	194846	194822
619	240O17_region_G3_4566_16_Forward_Primer	4455	4479
620	240O17_region_G3_4566_16_Reverse_Primer	4722	4696
621	240O17_region_G3_96209_14_Forward_Primer	96119	96143
622	240O17_region_G3_96209_14_Reverse_Primer	96392	96368
623	240O17_region_G3_248715_17_Forward_Primer	248633	248657
624	240O17_region_G3_248715_17_Reverse_Primer	248906	248882
625	240O17_region_G3_71410_40_Forward_Primer	71357	71379
626	240O17_region_G3_71410_40_Reverse_Primer	71636	71611
627	240O17_region_G3_226519_13_Forward_Primer	226315	226339
628	240O17_region_G3_226519_13_Reverse_Primer	226598	226574
629	240O17_region_G3_11282_19_Forward_Primer	11217	11242
630	240O17_region_G3_11282_19_Reverse_Primer	11501	11477
631	240O17_region_G3_170504_12_Forward_Primer	170409	170433
632	240O17_region_G3_170504_12_Reverse_Primer	170694	170671
633	240O17_region_G3_40864_14_Forward_Primer	40652	40678
634	240O17_region_G3_40864_14_Reverse_Primer	40938	40912
635	240O17_region_G3_13529_14_Forward_Primer	13332	13356
636	240O17_region_G3_13529_14_Reverse_Primer	13622	13598
637	240O17_region_G3_22858_14_Forward_Primer	22675	22699
638	240O17_region_G3_22858_14_Reverse_Primer	22966	22942
639	240O17_region_G3_309211_13_Forward_Primer	309092	309118
640	240O17_region_G3_309211_13_Reverse_Primer	309383	309358
641	240O17_region_G3_55568_26_Forward_Primer	55375	55399
642	240O17_region_G3_55568_26_Reverse_Primer	55667	55642
643	240O17_region_G3_73238_16_Forward_Primer	73043	73069

Seq Num	Seq ID	location of primer on config start	location of primer on config end
644	240017_region_G3_73238_16_Reverse_Primer	73342	73318
645	240017_region_G3_52488_19_Forward_Primer	52413	52437
646	240017_region_G3_52488_19_Reverse_Primer	52712	52688
647	318013_region_A3_471518_14_Forward_Primer_Seq	471464	471488
648	318013_region_A3_471518_14_Reverse_Primer_Seq	471567	471541
649	318013_region_A3_231599_23_Forward_Primer_Seq	231568	231592
650	318013_region_A3_231599_23_Reverse_Primer_Seq	231672	231651
651	318013_region_A3_375912_13_Forward_Primer_Seq	375845	375865
652	318013_region_A3_375912_13_Reverse_Primer_Seq	375954	375932
653	318013_region_A3_180013_12_Forward_Primer_Seq	179951	179974
654	318013_region_A3_180013_12_Reverse_Primer_Seq	180060	180038
655	318013_region_A3_171606_14_Forward_Primer_Seq	171545	171569
656	318013_region_A3_171606_14_Reverse_Primer_Seq	171657	171633
657	318013_region_A3_416256_13_Forward_Primer_Seq	416180	416203
658	318013_region_A3_416256_13_Reverse_Primer_Seq	416293	416269
659	318013_region_A3_231395_15_Forward_Primer_Seq	231339	231363
660	318013_region_A3_231395_15_Reverse_Primer_Seq	231461	231438
661	318013_region_A3_5502_47Forward_Primer_Seq	5461	5485
662	318013_region_A3_5502_47_Reverse_Primer_Seq	5585	5561
663	318013_region_A3_93061_14_Forward_Primer_Seq	92988	93012
664	318013_region_A3_93061_14_Reverse_Primer_Seq	93112	93090
665	318013_region_A3_111684_19_Forward_Primer_Seq	111646	111670
666	318013_region_A3_111684_19_Reverse_Primer_Seq	111772	111748
667	318013_region_A3_69328_14_Forward_Primer_Seq	69246	69269
668	318013_region_A3_69328_14_Reverse_Primer_Seq	69373	69349
669	318013_region_A3_36529_17_Forward_Primer_Seq	36488	36512
670	318013_region_A3_36529_17_Reverse_Primer_Seq	36617	36593
671	318013_region_A3_139128_12_Forward_Primer_Seq	139043	139067
672	318013_region_A3_139128_12_Reverse_Primer_Seq	139174	139150
673	318013_region_A3_495674_13_Forward_Primer_Seq	495592	495616
674	318013_region_A3_495674_13_Reverse_Primer_Seq	495723	495699
675	318013_region_A3_187577_13_Forward_Primer_Seq	187482	187506
676	318013_region_A3_187577_13_Reverse_Primer_Seq	187613	187590
677	318013_region_A3_453036_14_Forward_Primer_Seq	452999	453023
678	318013_region_A3_453036_14_Reverse_Primer_Seq	453132	453108
679	318013_region_A3_374041_13_Forward_Primer_Seq	373964	373988
680	318013_region_A3_374041_13_Reverse_Primer_Seq	374097	374073
681	318013_region_A3_3412_11_Forward_Primer_Seq	3319	3341
682	318013_region_A3_3412_11_Reverse_Primer_Seq	3454	3430
683	318013_region_A3_276495_28_Forward_Primer_Seq	276462	276485
684	318013_region_A3_276495_28_Reverse_Primer_Seq	276598	276574
685	318013_region_A3_151839_17_Forward_Primer_Seq	151744	151768
686	318013_region_A3_151839_17_Reverse_Primer_Seq	151882	151858
687	318013_region_A3_292912_12_Forward_Primer_Seq	292875	292899
688	318013_region_A3_292912_12_Reverse_Primer_Seq	293014	292990
689	318013_region_A3_104560_12_Forward_Primer_Seq	104464	104488
690	318013_region_A3_104560_12_Reverse_Primer_Seq	104604	104580
691	318013_region_A3_65193_11_Forward_Primer_Seq	65155	65179
692	318013_region_A3_65193_11_Reverse_Primer_Seq	65295	65271
693	318013_region_A3_110573_70_Forward_Primer_Seq	110533	110559
694	318013_region_A3_110573_70_Reverse_Primer_Seq	110674	110648



Seq Num	Seq ID	location of primer on config start	location of primer on config end
695	318O13_region_A3_65117_12_Forward_Primer_Seq	65034	65058
696	318O13_region_A3_65117_12_Reverse_Primer_Seq	65177	65153
697	318O13_region_A3_490837_16_Forward_Primer_Seq	490762	490786
698	318O13_region_A3_490837_16_Reverse_Primer_Seq	490905	490881
699	318O13_region_A3_107448_11_Forward_Primer_Seq	107385	107411
700	318O13_region_A3_107448_11_Reverse_Primer_Seq	107529	107505
701	318O13_region_A3_331_23_Forward_Primer_Seq	276	301
702	318O13_region_A3_331_23_Reverse_Primer_Seq	421	397
703	318O13_region_A3_193470_13_Forward_Primer_Seq	193444	193468
704	318O13_region_A3_193470_13_Reverse_Primer_Seq	193589	193565
705	318O13_region_A3_183305_14_Forward_Primer_Seq	183239	183263
706	318O13_region_A3_183305_14_Reverse_Primer_Seq	183384	183360
707	318O13_region_A3_55050_14_Forward_Primer_Seq	54998	55022
708	318O13_region_A3_55050_14_Reverse_Primer_Seq	55144	55120
709	318O13_region_A3_224693_21_Forward_Primer_Seq	224656	224682
710	318O13_region_A3_224693_21_Reverse_Primer_Seq	224803	224779
711	318O13_region_A3_207216_12_Forward_Primer_Seq	207152	207176
712	318O13_region_A3_207216_12_Reverse_Primer_Seq	207299	207276
713	318O13_region_A3_4654_22_Forward_Primer_Seq	4612	4636
714	318O13_region_A3_4654_22_Reverse_Primer_Seq	4760	4736
715	318O13_region_A3_408959_13_Forward_Primer_Seq	408918	408942
716	318O13_region_A3_408959_13_Reverse_Primer_Seq	409066	409042
717	318O13_region_A3_132288_22_Forward_Primer_Seq	132192	132216
718	318O13_region_A3_132288_22_Reverse_Primer_Seq	132340	132316
719	318O13_region_A3_292822_20_Forward_Primer_Seq	292747	292771
720	318O13_region_A3_292822_20_Reverse_Primer_Seq	292895	292871
721	318O13_region_A3_311076_12_Forward_Primer_Seq	311027	311051
722	318O13_region_A3_311076_12_Reverse_Primer_Seq	311175	311152
723	318O13_region_A3_509623_13_Forward_Primer_Seq	509584	509608
724	318O13_region_A3_509623_13_Reverse_Primer_Seq	509732	509708
725	318O13_region_A3_190404_14_Forward_Primer_Seq	190358	190382
726	318O13_region_A3_190404_14_Reverse_Primer_Seq	190506	190482
727	318O13_region_A3_164916_15_Forward_Primer_Seq	164808	164832
728	318O13_region_A3_164916_15_Reverse_Primer_Seq	164957	164933
729	318O13_region_A3_21028_13_Forward_Primer_Seq	21001	21026
730	318O13_region_A3_21028_13_Reverse_Primer_Seq	21150	21126
731	318O13_region_A3_208012_17_Forward_Primer_Seq	207955	207979
732	318O13_region_A3_208012_17_Reverse_Primer_Seq	208104	208080
733	318O13_region_A3_484089_14_Forward_Primer_Seq	484036	484060
734	318O13_region_A3_484089_14_Reverse_Primer_Seq	484185	484161
735	318O13_region_A3_332780_17_Forward_Primer_Seq	332723	332747
736	318O13_region_A3_332780_17_Reverse_Primer_Seq	332872	332853
737	318O13_region_A3_480137_37_Forward_Primer_Seq	480059	480084
738	318O13_region_A3_480137_37_Reverse_Primer_Seq	480208	480182
739	318O13_region_A3_441056_14_Forward_Primer_Seq	441011	441035
740	318O13_region_A3_441056_14_Reverse_Primer_Seq	441161	441138
741	318O13_region_A3_77486_11_Forward_Primer_Seq	77447	77471
742	318O13_region_A3_77486_11_Reverse_Primer_Seq	77597	77573
743	318O13_region_A3_272468_11_Forward_Primer_Seq	272423	272447
744	318O13_region_A3_272468_11_Reverse_Primer_Seq	272573	272549
745	318O13_region_A3_425319_17_Forward_Primer_Seq	425233	425257

Seq Num	Seq ID	location of primer on config start	location of primer on config end
746	318O13_region_A3_425319_17_Reverse_Primer_Seq	425383	425359
747	318O13_region_A3_413879_31_Forward_Primer_Seq	413835	413859
748	318O13_region_A3_413879_31_Reverse_Primer_Seq	413985	413961
749	318O13_region_A3_80477_64_Forward_Primer_Seq	80440	80464
750	318O13_region_A3_80477_64_Reverse_Primer_Seq	80591	80567
751	318O13_region_A3_277272_50_Forward_Primer_Seq	277213	277237
752	318O13_region_A3_277272_50_Reverse_Primer_Seq	277364	277340
753	318O13_region_A3_509642_13_Forward_Primer_Seq	509604	509628
754	318O13_region_A3_509642_13_Reverse_Primer_Seq	509755	509731
755	318O13_region_A3_321771_14_Forward_Primer_Seq	321663	321687
756	318O13_region_A3_321771_14_Reverse_Primer_Seq	321815	321791
757	318O13_region_A3_26788_12_Forward_Primer_Seq	26734	26758
758	318O13_region_A3_26788_12_Reverse_Primer_Seq	26886	26862
759	318O13_region_A3_262706_16_Forward_Primer_Seq	262649	262673
760	318O13_region_A3_262706_16_Reverse_Primer_Seq	262802	262778
761	318O13_region_A3_243928_16_Forward_Primer_Seq	243891	243915
762	318O13_region_A3_243928_16_Reverse_Primer_Seq	244044	244020
763	318O13_region_A3_23246_148_Forward_Primer_Seq	23215	23239
764	318O13_region_A3_23246_148_Reverse_Primer_Seq	23368	23344
765	318O13_region_A3_165406_12_Forward_Primer_Seq	165367	165391
766	318O13_region_A3_165406_12_Reverse_Primer_Seq	165521	165497
767	318O13_region_A3_486294_14_Forward_Primer_Seq	486208	486232
768	318O13_region_A3_486294_14_Reverse_Primer_Seq	486362	486338
769	318O13_region_A3_46754_12_Forward_Primer_Seq	46661	46685
770	318O13_region_A3_46754_12_Reverse_Primer_Seq	46816	46792
771	318O13_region_A3_381116_15_Forward_Primer_Seq	381080	381104
772	318O13_region_A3_381116_15_Reverse_Primer_Seq	381235	381211
773	318O13_region_A3_350369_11_Forward_Primer_Seq	350295	350319
774	318O13_region_A3_350369_11_Reverse_Primer_Seq	350450	350426
775	318O13_region_A3_138841_13_Forward_Primer_Seq	138795	138819
776	318O13_region_A3_138841_13_Reverse_Primer_Seq	138950	138926
777	318O13_region_A3_12158_142_Forward_Primer_Seq	12117	12141
778	318O13_region_A3_12158_142_Reverse_Primer_Seq	12272	12248
779	318O13_region_A3_315368_13_Forward_Primer_Seq	315310	315334
780	318O13_region_A3_315368_13_Reverse_Primer_Seq	315465	315441
781	318O13_region_A3_307549_13_Forward_Primer_Seq	307464	307488
782	318O13_region_A3_307549_13_Reverse_Primer_Seq	307619	307595
783	318O13_region_A3_159857_14_Forward_Primer_Seq	159772	159796
784	318O13_region_A3_159857_14_Reverse_Primer_Seq	159928	159904
785	318O13_region_A3_140551_15_Forward_Primer_Seq	140454	140478
786	318O13_region_A3_140551_15_Reverse_Primer_Seq	140610	140586
787	318O13_region_A3_279869_11_Forward_Primer_Seq	279797	279821
788	318O13_region_A3_279869_11_Reverse_Primer_Seq	279953	279929
789	318O13_region_A3_78292_35_Forward_Primer_Seq	78265	78291
790	318O13_region_A3_78292_35_Reverse_Primer_Seq	78422	78397
791	318O13_region_A3_185019_12_Forward_Primer_Seq	184953	184977
792	318O13_region_A3_185019_12_Reverse_Primer_Seq	185111	185087
793	318O13_region_A3_409164_13_Forward_Primer_Seq	409082	409106
794	318O13_region_A3_409164_13_Reverse_Primer_Seq	409240	409219
795	318O13_region_A3_75392_14_Forward_Primer_Seq	75287	75311
796	318O13_region_A3_75392_14_Reverse_Primer_Seq	75445	75421

Seq Num	Seq ID	location of primer on contig start	location of primer on contig end
797	318O13_region_A3_231320_12_Forward_Primer_Seq	231269	231293
798	318O13_region_A3_231320_12_Reverse_Primer_Seq	231429	231405
799	318O13_region_A3_381102_14_Forward_Primer_Seq	381041	381064
800	318O13_region_A3_381102_14_Reverse_Primer_Seq	381201	381176
801	318O13_region_A3_491826_15_Forward_Primer_Seq	491753	491777
802	318O13_region_A3_491826_15_Reverse_Primer_Seq	491914	491891
803	318O13_region_A3_56365_21_Forward_Primer_Seq	56336	56360
804	318O13_region_A3_56365_21_Reverse_Primer_Seq	56497	56473
805	318O13_region_A3_372628_15_Forward_Primer_Seq	372554	372578
806	318O13_region_A3_372628_15_Reverse_Primer_Seq	372715	372691
807	318O13_region_A3_217037_11_Forward_Primer_Seq	216919	216943
808	318O13_region_A3_217037_11_Reverse_Primer_Seq	217081	217057
809	318O13_region_A3_302609_11_Forward_Primer_Seq	302575	302599
810	318O13_region_A3_302609_11_Reverse_Primer_Seq	302737	302713
811	318O13_region_A3_341804_11_Forward_Primer_Seq	341686	341710
812	318O13_region_A3_341804_11_Reverse_Primer_Seq	341848	341824
807	318O13_region_A3_217037_11_Forward_Primer_Seq	216919	216943
808	318O13_region_A3_217037_11_Reverse_Primer_Seq	217081	217057
813	318O13_region_A3_264929_68_Forward_Primer_Seq	264862	264886
814	318O13_region_A3_264929_68_Reverse_Primer_Seq	265024	265000
815	318O13_region_A3_55499_12_Forward_Primer_Seq	55400	55424
816	318O13_region_A3_55499_12_Reverse_Primer_Seq	55563	55539
817	318O13_region_A3_295634_14_Forward_Primer_Seq	295538	295562
818	318O13_region_A3_295634_14_Reverse_Primer_Seq	295702	295677
819	318O13_region_A3_269358_15_Forward_Primer_Seq	269242	269266
820	318O13_region_A3_269358_15_Reverse_Primer_Seq	269406	269382
821	318O13_region_A3_457009_24_Forward_Primer_Seq	456924	456948
822	318O13_region_A3_457009_24_Reverse_Primer_Seq	457088	457064
823	318O13_region_A3_176598_14_Forward_Primer_Seq	176554	176578
824	318O13_region_A3_176598_14_Reverse_Primer_Seq	176718	176694
825	318O13_region_A3_278266_12_Forward_Primer_Seq	278210	278234
826	318O13_region_A3_278266_12_Reverse_Primer_Seq	278376	278350
827	318O13_region_A3_391810_12_Forward_Primer_Seq	391683	391707
828	318O13_region_A3_391810_12_Reverse_Primer_Seq	391851	391826
829	318O13_region_A3_269485_15_Forward_Primer_Seq	269383	269407
830	318O13_region_A3_269485_15_Reverse_Primer_Seq	269551	269527
831	318O13_region_A3_359247_17_Forward_Primer_Seq	359218	359243
832	318O13_region_A3_359247_17_Reverse_Primer_Seq	359386	359362
833	318O13_region_A3_315094_13_Forward_Primer_Seq	314976	315002
834	318O13_region_A3_315094_13_Reverse_Primer_Seq	315145	315120
835	318O13_region_A3_307823_13_Forward_Primer_Seq	307784	307809
836	318O13_region_A3_307823_13_Reverse_Primer_Seq	307953	307927
837	318O13_region_A3_248588_15_Forward_Primer_Seq	248540	248564
838	318O13_region_A3_248588_15_Reverse_Primer_Seq	248709	248684
839	318O13_region_A3_252426_85_Forward_Primer_Seq	252398	252423
840	318O13_region_A3_252426_85_Reverse_Primer_Seq	252568	252543
841	318O13_region_A3_513314_16_Forward_Primer_Seq	513209	513233
842	318O13_region_A3_513314_16_Reverse_Primer_Seq	513379	513355
843	318O13_region_A3_68183_14_Forward_Primer_Seq	68108	68132
844	318O13_region_A3_68183_14_Reverse_Primer_Seq	68279	68255
845	318O13_region_A3_471191_13_Forward_Primer_Seq	471059	471083

Seq Num	Seq ID	location of primer on contig start	location of primer on contig end
846	318013_region_A3_471191_13 Reverse_Primer_Seq	471231	471206
847	318013_region_A3_163547_18 Forward_Primer_Seq	163459	163483
848	318013_region_A3_163547_18 Reverse_Primer_Seq	163632	163608
849	318013_region_A3_417867_15 Forward_Primer_Seq	417839	417863
850	318013_region_A3_417867_15 Reverse_Primer_Seq	418014	417990
851	318013_region_A3_332465_14 Forward_Primer_Seq	332346	332370
852	318013_region_A3_332465_14 Reverse_Primer_Seq	332523	332499
853	318013_region_A3_207697_14 Forward_Primer_Seq	207578	207602
854	318013_region_A3_207697_14 Reverse_Primer_Seq	207755	207731
855	318013_region_A3_277229_43 Forward_Primer_Seq	277186	277210
856	318013_region_A3_277229_43 Reverse_Primer_Seq	277364	277340
857	318013_region_A3_36366_11 Forward_Primer_Seq	36323	36345
858	318013_region_A3_36366_11 Reverse_Primer_Seq	36501	36477
859	318013_region_A3_91970_12 Forward_Primer_Seq	91938	91962
860	318013_region_A3_91970_12 Reverse_Primer_Seq	92116	92091
861	318013_region_A3_211533_11 Forward_Primer_Seq	211406	211430
862	318013_region_A3_211533_11 Reverse_Primer_Seq	211585	211561
863	318013_region_A3_336301_11 Forward_Primer_Seq	336174	336198
864	318013_region_A3_336301_11 Reverse_Primer_Seq	336353	336329
865	318013_region_A3_441603_14 Forward_Primer_Seq	441539	441563
866	318013_region_A3_441603_14 Reverse_Primer_Seq	441718	441694
867	318013_region_A3_468354_15 Forward_Primer_Seq	468263	468287
868	318013_region_A3_468354_15 Reverse_Primer_Seq	468442	468418
869	318013_region_A3_188983_18 Forward_Primer_Seq	188855	188879
870	318013_region_A3_188983_18 Reverse_Primer_Seq	189035	189009
871	318013_region_A3_115502_17 Forward_Primer_Seq	115469	115493
872	318013_region_A3_115502_17 Reverse_Primer_Seq	115649	115625
873	318013_region_A3_163006_13 Forward_Primer_Seq	162972	162996
874	318013_region_A3_163006_13 Reverse_Primer_Seq	163153	163129
875	318013_region_A3_119283_14 Forward_Primer_Seq	119199	119224
876	318013_region_A3_119283_14 Reverse_Primer_Seq	119381	119357
877	318013_region_A3_491126_11 Forward_Primer_Seq	491062	491086
878	318013_region_A3_491126_11 Reverse_Primer_Seq	491244	491220
879	318013_region_A3_99512_21 Forward_Primer_Seq	99398	99422
880	318013_region_A3_99512_21 Reverse_Primer_Seq	99581	99557
881	318013_region_A3_280291_17 Forward_Primer_Seq	280201	280226
882	318013_region_A3_280291_17 Reverse_Primer_Seq	280385	280361
883	318013_region_A3_138443_19 Forward_Primer_Seq	138304	138329
884	318013_region_A3_138443_19 Reverse_Primer_Seq	138488	138465
885	318013_region_A3_115973_14 Forward_Primer_Seq	115832	115856
886	318013_region_A3_115973_14 Reverse_Primer_Seq	116016	115992
887	318013_region_A3_329977_14 Forward_Primer_Seq	329864	329889
888	318013_region_A3_329977_14 Reverse_Primer_Seq	330050	330026
889	318013_region_A3_205203_14 Forward_Primer_Seq	205090	205115
890	318013_region_A3_205203_14 Reverse_Primer_Seq	205276	205252
891	318013_region_A3_153114_12 Forward_Primer_Seq	152969	152993
892	318013_region_A3_153114_12 Reverse_Primer_Seq	153156	153132
893	318013_region_A3_34581_13 Forward_Primer_Seq	34523	34547
894	318013_region_A3_34581_13 Reverse_Primer_Seq	34712	34688
895	318013_region_A3_292577_19 Forward_Primer_Seq	292549	292573
896	318013_region_A3_292577_19 Reverse_Primer_Seq	292739	292715

Seq Num	Seq ID	location of primer on contig start	location of primer on contig end
897	318O13_region_A3_445391_20_Forward_Primer_Seq	445356	445382
898	318O13_region_A3_445391_20_Reverse_Primer_Seq	445547	445523
899	318O13_region_A3_350540_17_Forward_Primer_Seq	350421	350445
900	318O13_region_A3_350540_17_Reverse_Primer_Seq	350612	350588
901	318O13_region_A3_453879_15_Forward_Primer_Seq	453725	453750
902	318O13_region_A3_453879_15_Reverse_Primer_Seq	453918	453894
903	318O13_region_A3_201246_13_Forward_Primer_Seq	201128	201153
904	318O13_region_A3_201246_13_Reverse_Primer_Seq	201321	201297
905	318O13_region_A3_326020_13_Forward_Primer_Seq	325902	325927
906	318O13_region_A3_326020_13_Reverse_Primer_Seq	326095	326071
907	318O13_region_A3_503801_14_Forward_Primer_Seq	503656	503680
908	318O13_region_A3_503801_14_Reverse_Primer_Seq	503849	503823
909	318O13_region_A3_302400_52_Forward_Primer_Seq	302283	302307
910	318O13_region_A3_302400_52_Reverse_Primer_Seq	302481	302456
911	318O13_region_A3_448857_15_Forward_Primer_Seq	448748	448772
912	318O13_region_A3_448857_15_Reverse_Primer_Seq	448947	448924
913	318O13_region_A3_48364_14_Forward_Primer_Seq	48232	48256
914	318O13_region_A3_48364_14_Reverse_Primer_Seq	48435	48412
915	318O13_region_A3_251804_48_Forward_Primer_Seq	251738	251762
916	318O13_region_A3_251804_48_Reverse_Primer_Seq	251942	251918
917	318O13_region_A3_382583_13_Forward_Primer_Seq	382549	382574
918	318O13_region_A3_382583_13_Reverse_Primer_Seq	382753	382728
919	318O13_region_A3_124737_14_Forward_Primer_Seq	124641	124665
920	318O13_region_A3_124737_14_Reverse_Primer_Seq	124846	124822
921	318O13_region_A3_124766_13_Forward_Primer_Seq	124641	124665
922	318O13_region_A3_124766_13_Reverse_Primer_Seq	124846	124822
923	318O13_region_A3_461351_16_Forward_Primer_Seq	461218	461242
924	318O13_region_A3_461351_16_Reverse_Primer_Seq	461426	461402
925	318O13_region_A3_64953_19_Forward_Primer_Seq	64798	64823
926	318O13_region_A3_64953_19_Reverse_Primer_Seq	65011	64987
927	318O13_region_A3_366586_13_Forward_Primer_Seq	366508	366532
928	318O13_region_A3_366586_13_Reverse_Primer_Seq	366722	366698
929	318O13_region_A3_46190_15_Forward_Primer_Seq	46012	46037
930	318O13_region_A3_46190_15_Reverse_Primer_Seq	46228	46205
931	318O13_region_A3_81016_11_Forward_Primer_Seq	80927	80952
932	318O13_region_A3_81016_11_Reverse_Primer_Seq	81146	81122
933	318O13_region_A3_134426_14_Forward_Primer_Seq	134253	134277
934	318O13_region_A3_134426_14_Reverse_Primer_Seq	134474	134449
935	318O13_region_A3_292724_14_Forward_Primer_Seq	292549	292573
936	318O13_region_A3_292724_14_Reverse_Primer_Seq	292771	292747
937	318O13_region_A3_187096_17_Forward_Primer_Seq	187058	187082
938	318O13_region_A3_187096_17_Reverse_Primer_Seq	187282	187257
939	318O13_region_A3_381693_13_Forward_Primer_Seq	381658	381683
940	318O13_region_A3_381693_13_Reverse_Primer_Seq	381885	381863
941	318O13_region_A3_361286_33_Forward_Primer_Seq	361173	361197
942	318O13_region_A3_361286_33_Reverse_Primer_Seq	361401	361376
943	318O13_region_A3_482668_14_Forward_Primer_Seq	482592	482616
944	318O13_region_A3_482668_14_Reverse_Primer_Seq	482821	482796
945	318O13_region_A3_128002_12_Forward_Primer_Seq	127882	127906
946	318O13_region_A3_128002_12_Reverse_Primer_Seq	128112	128087
947	318O13_region_A3_499270_14_Forward_Primer_Seq	499184	499208

Seq Num	Seq ID	location of primer on config start	location of primer on config end
948	318013_region_A3_499270_14_Reverse_Primer_Seq	499422	499398
949	318013_region_A3_231650_12_Forward_Primer_Seq	231568	231592
950	318013_region_A3_231650_12_Reverse_Primer_Seq	231809	231788
951	318013_region_A3_199851_13_Forward_Primer_Seq	199762	199786
952	318013_region_A3_199851_13_Reverse_Primer_Seq	200012	199988
953	318013_region_A3_324629_13_Forward_Primer_Seq	324540	324564
954	318013_region_A3_324629_13_Reverse_Primer_Seq	324790	324766
955	318013_region_A3_374190_19_Forward_Primer_Seq	374129	374152
956	318013_region_A3_374190_19_Reverse_Primer_Seq	374394	374370
957	318013_region_A3_460603_13_Forward_Primer_Seq	460450	460474
958	318013_region_A3_460603_13_Reverse_Primer_Seq	460715	460691
959	318013_region_A3_108681_14_Forward_Primer_Seq	108524	108548
960	318013_region_A3_108681_14_Reverse_Primer_Seq	108791	108768
961	318013_region_A3_459791_47_Forward_Primer_Seq	459639	459663
962	318013_region_A3_459791_47_Reverse_Primer_Seq	459907	459883
963	318013_region_A3_4257_20_Forward_Primer_Seq	4172	4196
964	318013_region_A3_4257_20_Reverse_Primer_Seq	4450	4425
965	318013_region_A3_238810_14_Forward_Primer_Seq	238563	238589
966	318013_region_A3_238810_14_Reverse_Primer_Seq	238850	238826
967	318013_region_A3_245817_14_Forward_Primer_Seq	245713	245738
968	318013_region_A3_245817_14_Reverse_Primer_Seq	246001	245977
969	318013_region_A3_245956_14_Forward_Primer_Seq	245713	245738
970	318013_region_A3_245956_14_Reverse_Primer_Seq	246001	245977
971	318013_region_A3_74148_14_Forward_Primer_Seq	74050	74075
972	318013_region_A3_74148_14_Reverse_Primer_Seq	74338	74314
973	318013_region_A3_74089_15_Forward_Primer_Seq	74050	74075
974	318013_region_A3_74089_15_Reverse_Primer_Seq	74338	74314
975	318013_region_A3_241686_12_Forward_Primer_Seq	241470	241494
976	318013_region_A3_241686_12_Reverse_Primer_Seq	241765	241741
977	318013_region_A3_47476_12_Forward_Primer_Seq	47280	47304
978	318013_region_A3_47476_12_Reverse_Primer_Seq	47577	47554
979	318013_region_A3_164550_12_Forward_Primer_Seq	164323	164347
980	318013_region_A3_164550_12_Reverse_Primer_Seq	164621	164598
981	318013_region_A3_101255_15_Forward_Primer_Seq	101119	101144
982	318013_region_A3_101255_15_Reverse_Primer_Seq	101418	101392
983	515002_region_G2_16189_11_Forward_Primer	16144	16168
984	515002_region_G2_16189_11_Reverse_Primer	16244	16220
985	515002_region_G2_71925_13_Forward_Primer	71880	71905
986	515002_region_G2_71925_13_Reverse_Primer	71987	71963
987	515002_region_G2_4707_12_Forward_Primer	4660	4684
988	515002_region_G2_4707_12_Reverse_Primer	4769	4743
989	515002_region_G2_118904_18_Forward_Primer	118847	118871
990	515002_region_G2_118904_18_Reverse_Primer	118957	118932
991	515002_region_G2_13655_17_Forward_Primer	13567	13592
992	515002_region_G2_13655_17_Reverse_Primer	13698	13673
993	515002_region_G2_53900_13_Forward_Primer	53843	53867
994	515002_region_G2_53900_13_Reverse_Primer	53985	53961
995	515002_region_G2_8079_14_Forward_Primer	8023	8047
996	515002_region_G2_8079_14_Reverse_Primer	8167	8143
997	515002_region_G2_9969_28_Forward_Primer	9917	9941
998	515002_region_G2_9969_28_Reverse_Primer	10062	10038

Seq Num	Seq ID	location of primer on config start	location of primer on config end
999	515002_region_G2_72308_77_Forward_Primer	72272	72298
1000	515002_region_G2_72308_77_Reverse_Primer	10062	10038
1001	515002_region_G2_99475_19_Forward_Primer	99408	99433
1002	515002_region_G2_99475_19_Reverse_Primer	99554	99530
1003	515002_region_G2_118615_18_Forward_Primer	118512	118535
1004	515002_region_G2_118615_18_Reverse_Primer	118658	118634
1005	515002_region_G2_119001_46_Forward_Primer	118931	118956
1006	515002_region_G2_119001_46_Reverse_Primer	119079	119055
1007	515002_region_G2_118958_43_Forward_Primer	118931	118956
1008	515002_region_G2_118958_43_Reverse_Primer	119079	119055
1009	515002_region_G2_17197_13_Forward_Primer	17128	17152
1010	515002_region_G2_17197_13_Reverse_Primer	17276	17252
1011	515002_region_G2_105163_29_Forward_Primer	105068	105092
1012	515002_region_G2_105163_29_Reverse_Primer	105217	105192
1013	515002_region_G2_111335_13_Forward_Primer	111308	111332
1014	515002_region_G2_111335_13_Reverse_Primer	111458	111434
1015	515002_region_G2_106396_13_Forward_Primer	106318	106342
1016	515002_region_G2_106396_13_Reverse_Primer	106469	106445
1017	515002_region_G2_59229_17_Forward_Primer	59203	59227
1018	515002_region_G2_59229_17_Reverse_Primer	59354	59330
1019	515002_region_G2_73795_20_Forward_Primer	73769	73793
1020	515002_region_G2_73795_20_Reverse_Primer	73921	73896
1021	515002_region_G2_85664_20_Forward_Primer	85586	85611
1022	515002_region_G2_85664_20_Reverse_Primer	85738	85714
1023	515002_region_G2_36921_17_Forward_Primer	36830	36854
1024	515002_region_G2_36921_17_Reverse_Primer	36983	36959
1025	515002_region_G2_124150_19_Forward_Primer	124073	124096
1026	515002_region_G2_124150_19_Reverse_Primer	124227	124203
1027	515002_region_G2_5089_14_Forward_Primer	4999	5024
1028	515002_region_G2_5089_14_Reverse_Primer	5156	5132
1029	515002_region_G2_58221_15_Forward_Primer	58197	58220
1030	515002_region_G2_58221_15_Reverse_Primer	58354	58330
1031	515002_region_G2_96139_14_Forward_Primer	96022	96046
1032	515002_region_G2_96139_14_Reverse_Primer	96182	96158
1033	515002_region_G2_70595_13_Forward_Primer	70472	70496
1034	515002_region_G2_70595_13_Reverse_Primer	70634	70608
1035	515002_region_G2_4340_15_Forward_Primer	4312	4337
1036	515002_region_G2_4340_15_Reverse_Primer	4477	4454
1037	515002_region_G2_90417_11_Forward_Primer	90335	90359
1038	515002_region_G2_90417_11_Reverse_Primer	90503	90479
1039	515002_region_G2_49711_17_Forward_Primer	49652	49676
1040	515002_region_G2_49711_17_Reverse_Primer	49820	49796
1041	515002_region_G2_63053_13_Forward_Primer	63005	63029
1042	515002_region_G2_63053_13_Reverse_Primer	63173	63148
1043	515002_region_G2_63076_14_Forward_Primer	63005	63029
1044	515002_region_G2_63076_14_Reverse_Primer	63173	63148
1045	515002_region_G2_44442_12_Forward_Primer	44335	44359
1046	515002_region_G2_44442_12_Reverse_Primer	44505	44481
1047	515002_region_G2_44422_19_Forward_Primer	44335	44359
1048	515002_region_G2_44422_19_Reverse_Primer	44505	44481
1049	515002_region_G2_44158_19_Forward_Primer	44075	44100

Seq Num	Seq ID	location of primer on contig start	location of primer on contig end
1050	515O02_region_G2_44158_19_Reverse_Primer	44252	44227
1051	515O02_region_G2_44141_17_Forward_Primer	44075	44100
1052	515O02_region_G2_44141_17_Reverse_Primer	44252	44227
1053	515O02_region_G2_90762_17_Forward_Primer	90637	90663
1054	515O02_region_G2_90762_17_Reverse_Primer	90814	90790
1055	515O02_region_G2_106241_14_Forward_Primer	106160	106184
1056	515O02_region_G2_106241_14_Reverse_Primer	106341	106317
1057	515O02_region_G2_109676_12_Forward_Primer	109609	109634
1058	515O02_region_G2_109676_12_Reverse_Primer	109793	109768
1059	515O02_region_G2_86242_14_Forward_Primer	86134	86158
1060	515O02_region_G2_86242_14_Reverse_Primer	86318	86293
1061	515O02_region_G2_83109_12_Forward_Primer	83017	83041
1062	515O02_region_G2_83109_12_Reverse_Primer	83202	83178
1063	515O02_region_G2_10461_15_Forward_Primer	10418	10442
1064	515O02_region_G2_10461_15_Reverse_Primer	10609	10585
1065	515O02_region_G2_67608_15_Forward_Primer	67552	67577
1066	515O02_region_G2_67608_15_Reverse_Primer	67745	67721
1067	515O02_region_G2_63275_46_Forward_Primer	63148	63173
1068	515O02_region_G2_63275_46_Reverse_Primer	63347	63323
1069	515O02_region_G2_62405_14_Forward_Primer	62374	62399
1070	515O02_region_G2_62405_14_Reverse_Primer	62576	62552
1071	515O02_region_G2_33563_12_Forward_Primer	33460	33484
1072	515O02_region_G2_33563_12_Reverse_Primer	33670	33646
1073	515O02_region_G2_33146_14_Forward_Primer	32949	32973
1074	515O02_region_G2_33146_14_Reverse_Primer	33191	33167
1075	515O02_region_G2_102179_29_Forward_Primer	102102	102126
1076	515O02_region_G2_102179_29_Reverse_Primer	102352	102327
1077	515O02_region_G2_2646_15_Forward_Primer	2553	2577
1078	515O02_region_G2_2646_15_Reverse_Primer	2809	2784
1079	515O02_region_G2_76652_24_Forward_Primer	76567	76591
1080	515O02_region_G2_76652_24_Reverse_Primer	76835	76812
1081	515O02_region_G2_66280_14_Forward_Primer	66052	66077
1082	515O02_region_G2_66280_14_Reverse_Primer	66334	66309
1083	515O02_region_G2_54768_13_Forward_Primer	54640	54666
1084	515O02_region_G2_54768_13_Reverse_Primer	54923	54899
1085	515O02_region_G2_62580_14_Forward_Primer	62552	62576
1086	515O02_region_G2_62580_14_Reverse_Primer	62840	62816
1087	515O02_region_G2_34598_55_Forward_Primer	34473	34497
1088	515O02_region_G2_34598_55_Reverse_Primer	34765	34739
1089	515O02_region_G2_77680_13_Forward_Primer	77444	77470
1090	515O02_region_G2_77680_13_Reverse_Primer	77741	77716
1091	515O02_region_G2_77693_12_Forward_Primer	77444	77470
1092	515O02_region_G2_77693_12_Reverse_Primer	77741	77716
1093	515O02_region_G2_97392_14_Forward_Primer	97255	97280
1094	515O02_region_G2_97392_14_Reverse_Primer	97554	97530
1095	515O02_region_G2_97359_15_Forward_Primer	97255	97280
1096	515O02_region_G2_97359_15_Reverse_Primer	97554	97530

Seq Num	Seq ID		
1120	consensusLRR		
1121	rhgILRR		



Seq Num	Seq ID		
1122	Rhg4LRR		

Seq Num	Seq ID	Primer location on 240O17_region_G3	
1123	240O17_region_G3_forward_1_b	45046-45072	

## DETAILED DESCRIPTION OF THE INVENTION

A) *rhg1*

The present invention provides a method for the production of a soybean plant having an *rhg1* SCN resistant allele comprising: (A) crossing a first soybean plant having an *rhg1* SCN resistant allele with a second soybean plant having an *rhg1* SCN sensitive allele to produce a segregating population; (B) screening the segregating population for a member having an *rhg1* SCN resistant allele with a first nucleic acid molecule capable of specifically hybridizing to linkage group G, wherein the first nucleic acid molecule specifically hybridizes to a second nucleic acid molecule that is linked to the *rhg1* SCN resistant allele; and, (C) selecting the member for further crossing and selection.

*rhg1* is located on linkage group G (Concibido *et al.*, *Crop Sci.* 36:1643-1650 (1996)). SCN resistant alleles of *rhg1* provide partial resistance to SCN races 1, 2, 3, 5, 6, and 14 (Concibido *et al.* (*Crop Sci.* 37:258-264 (1997)). Also, Webb (U.S. Patent 5,491,081) reports that a QTL on linkage group G (*rhg1*) provides partial resistance to SCN races 1, 2, 3, 5, and 14. *rhg1* and *Rhg4* provide complete or nearly complete resistance to SCN race 3 (U.S. Patent 5,491,081). While initially thought to be a recessive gene, *rhg1* classification as a recessive gene has been questioned.

Using bioinformatic approaches, the *rhg1* coding region is predicted to contain either four exons (*rhg1*, v.1)(coding coordinates 45163-45314, 45450-45509, 46941-48763, and 48975-49573 of SEQ ID NO: 2) or two exons (*rhg1*, v.2) (coding coordinates 46798-48763 and 48975-49573 of SEQ ID NO: 3). *rhg1*, v.1 encodes an 877 amino acid polypeptide. *rhg1*, v.2 encodes an 854 amino acid length polypeptide. *rhg1* codes for a Xa21-like receptor kinase (SEQ ID NOs: 1097, 1098, and 1100-1115) (Song, *et al.*, *Science* 270, 1804-1806 (1995)). *rhg1* has an extracellular leucine rich repeat (LRR) domain (*rhg1*, v.1, SEQ ID NO: 1097, residues 164-457; *rhg1*, v.2, SEQ ID NO: 1098,

residues 141-434), a transmembrane domain (*rhg1*, v.1, SEQ ID NO: 1097, residues 508-530; *rhg1*, v.2, SEQ ID NO: 1098, residues 33-51, and 485-507) and serine/threonine protein kinase (STK) domain (*rhg1*, v.1, SEQ ID NO: 1097, residues 578-869; *rhg1*, v.2, SEQ ID NO: 1098, residues 555-846). In a preferred embodiment, the LRR domain has multiple LRR repeats. In a more preferred embodiment, the LRR domain has 12 LRR repeats.

To identify proteins similar to the proteins encoded by *rhg1* candidates, database searches are performed using the predicted peptide sequences. The *rhg1* candidate shows similarity to CAA18124, which is the *Arabidopsis* putative receptor kinase (58.4% similarity and 35.8% identity, (CLUSTALW (default parameters), Thompson *et al.*, Nucleic Acids Res. 22:4673-4680 (1994)), GCG package, Genetics Computer Group, Madison, Wisconsin), and the apple leucine-rich receptor-like protein kinase (g3641252) (53.2% similarity and 31.5% identity, (CLUSTALW (default parameters))), which has both LLR and STK domains, showing conservation in both the LLR and STK domains. The predicted LRR extracellular domain shows similarity to the tomato resistance genes *Cf-2.1* (*Lycopersicon pimpinellifolium*) (66.9% similarity and 45.4% identity (CLUSTALW (default parameters))) and *Cf-2.2* (*Lycopersicon pimpinellifolium*) (66.9% similarity and 45.4% identity (CLUSTALW (default parameters))).

Figure 1 is an alignment of the LRR domain of the *rhg1* gene. A consensus sequence for the LRR is shown as the top row of the alignment. Each row of amino acids represents an LRR domain. The boxed region indicates the putative  $\beta$ -turn/ $\beta$ -sheet structural motif postulated to be involved in ligand binding (Jones *et al.*, *Adv. Bot. Res. Incorp. Adv. Plant Path.* 24:89-167 (1997)). The hydrophobic leucine residues are thought to project into the core of the protein while the flanking amino acids are thought to be solvent exposed where they may interact with the ligand (Kobe *et al.*, *Nature* 374:183-186 (1995)). Non-conservative changes in this region are thought to affect folding. An "x" represents an arbitrary amino acid while an "a" represents a hydrophobic residue (leucine, isoleucine, methionine, valine, or phenylalanine). Amino acid substitutions

between resistant and sensitive phenotypes are bordered by a double line. The amino acid substitution within the 302-325 region is a histidine/asparagine substitution, and the amino acid substitution within the 422-445 region is a phenylalanine/serine substitution.

As used herein, a naturally occurring *rhg1* allele is any allele that encodes for a protein having an extracellular LRR, a transmembrane domain, and STK domain where the naturally occurring allele is present on linkage group G and where certain *rhg1* alleles, but not all *rhg1* alleles, are capable of providing or contributing to resistance or partial resistance to a race of SCN. It is understood that such an allele can, using for example methods disclosed herein, be manipulated so that the nucleic acid molecule encoding the protein is no longer present on linkage group G. It is also understood that such an allele can, using for example methods disclosed herein, be manipulated so that the nucleic acid molecule sequence is altered.

As used herein, an *rhg1* SCN resistant allele is any *rhg1* allele where that allele alone or in combination with other SCN resistant alleles present in the plant, such as an *Rhg4* SCN resistant allele, provides resistance to a race of SCN, and that resistance is due, at least in part, to the genetic contribution of the *rhg1* allele.

SCN resistance or partial resistance is determined by a comparison of the plant in question with a known SCN sensitive host, Lee 74, according to the method set forth in Schmitt, *J. Nematol.* 20:392-395 (1988). As used herein, resistance to a particular race of SCN is defined as having less than 10% of cyst development relative to the SCN sensitive host Lee 74. Moreover, as used herein, partial resistance to a particular race of SCN is defined as having more than 10% but less than 75% of cyst development relative to the SCN sensitive host Lee 74.

Any soybean plant having an *rhg1* SCN resistant allele can be used in conjunction with the present invention. Soybeans with known *rhg1* SCN resistant alleles can be used. Such soybeans include but are not limited to PI548402 (Peking), PI200499, A2869, Jack, A2069, PI209332 (No:4), PI404166 (Krasnoarmejkaja), PI404198 (Sun huan do), PI437654 (Er-hej-jan), PI438489 (Chiquita), PI507354 (Tokei 421), PI548655 (Forrest),

PI548988 (Pickett), PI84751, PI437654, PI40792, Pyramid, Nathan, AG2201, A3469, AG3901, A3904, AG4301, AG4401, AG4501, AG4601, PION9492, PI88788, Dyer, Custer, Manokin, and Doles. In a preferred aspect, the soybean plant having an *rhg1* SCN resistant allele is an *rhg1* haplotype 2 allele. Examples of soybeans with an *rhg1* haplotype 2 allele are PI548402 (Peking), PI404166 (Krasnoarmejkaja), PI404198 (Sun huan do), PI437654 (Er-hej-jan), PI438489 (Chiquita), PI507354 (Tokei 421), PI548655 (Forrest), PI548988 (Pickett), PI84751, PI437654, and PI40792. In addition, using the methods or agents of the present invention, soybeans and wild relative of soybean such as *Glycine soja* can be screened for the presence of *rhg1* SCN resistant alleles.

Any soybean plant having an *rhg1* SCN sensitive allele can be used in conjunction with the present invention. Such soybeans include A3244, A2833, AG3001, Williams, Will, A2704, Noir, DK23-51, Lee 74, Essex, Minsoy, A1923, and Hutcheson. In a preferred aspect, the soybean plant having an *rhg1* SCN sensitive allele is an *rhg1* A3244 allele. In addition, using the methods or agents of the present invention, soybeans and wild relatives of soybean such as *Glycine soja* can be screened for the presence of *rhg1* SCN sensitive alleles.

Table 2, below, is a table showing single nucleotide polymorphisms (SNPs) and insertions/deletions (INDEL) sites for eight haplotype sequences of *rhg1*.

TABLE 2

Identification			Base number of contig 240017_region_G3 of reference line A3244															
Hap	PI#	Line	Ph	45173	45309	45400	45416	45439	45611	45916	45958	46049	46113	46227	46703	47057	47140	47208
1	-	A3244	S	G	G	A	T	A	A	A	A	C	A	d1	0	T	C	C
2	PI548402	Peking	R	G	A	C	C	T	A	G	A	T	G	0	d2	C	C	C
3	PI423871	Toyosazu	-	G	A	A	T	A	A	G	A	T	G	0	0	T	C	C
4	PI518672	Willi	S	G	G	A	T	A	A	A	A	C	A	d1	0	T	A	T
5	-	A2704	S	G	G	A	T	A	A	A	A	C	A	d1	0	T	A	T
6	PI290136	Nolr	S	A	A	A	C	T	G	A	T	T	A	0	d14	T	C	C
7	PI548658	Lee 74	S	A	A	A	C	T	G	A	T	T	A	0	d14	T	C	C
8	PI200499	-	R	G	A	A	C	A	A	A	A	T	A	0	d14	T	C	C
N/A	PI548667	Essex	S	A	A	A	C	T	G	A	T	T	A	0	d14	T	C	C
N/A	PI548389	Minsoy	S	G	G	A	T	A	A	A	A	C	A	d1	0	T	A	T
N/A	PI360843	Oshima.	-	-	-	-	-	-	-	-	-	-	-	-	0	T	A	T
N/A	-	A2869	R	-	-	-	-	-	-	-	-	-	-	0	d14	T	C	C
N/A	PI540556	Jack	R	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
N/A	-	A2069	R	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
N/A	PI209832	No.4	R	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

TABLE 2, continued

Identification			Base number of contig 240017_region_G3 of reference line A3244															
Hap	PI#	Line	Ph	47571	47617	47796	47856	47937	48012	48060	48073	48135	48279	48413	48681	48881	49012	49316
1	-	A3244	S	G	C	A	T	T	T	C	C	A	C	G	A	0	A	T
2	PI548402	Peking	R	G	C	C	C	C	T	C	C	G	C	G	G	d19	G	T
3	PI423871	Toyosuzu	-	G	C	C	C	C	T	C	C	G	C	G	A	0	A	T
4	PI518672	Willi	S	G	C	A	T	T	T	C	C	A	C	G	A	0	A	T
5	-	A2704	S	G	C	A	T	T	C	T	T	G	T	C	-	0	G	C
6	PI290136	Nolr	S	A	A	C	C	C	C	T	T	G	T	C	G	0	G	C
7	PI548658	Lee 74	S	G	A	C	C	C	C	T	T	G	T	C	G	0	G	C
8	PI200499	-	R	G	A	C	C	C	C	T	T	G	T	C	G	0	G	C
N/A	PI548667	Essex	S	G	A	C	C	C	C	T	T	G	T	C	A/G	0	G	C
N/A	PI548389	Minsoy	S	G	C	A	T	T	C/T	C/T	C/T	A	C	G	A	0	A	T
N/A	PI360843	Oshimas.	-	G	C	A	T/C	T/C	T	C	C	A/G	C	G	A	0	A	T
N/A	-	A2889	R	G	A	C	C	C	C	T	T	G	T	C	G	0	G	C
N/A	PI540556	Jack	R	-	-	C	C	C	C	T	T	G	T	C	G	0	G	C
N/A	-	A2069	R	-	-	C	T/C	T/C	C	T	T	A	T	C	G	0	G	C
N/A	PI208332	No.4	R	-	-	C	C	C	C	T	T	A/G	T	C	G	0	G	C

In Table 2, discrete haplotypes are designated 1 through 8. N/A refers to a haplotype that is not characterized. The Plant Introduction classification number is indicated in the "PI#" column. A dash indicates that no PI number is known or assigned for the line under investigation. The line from which the sequences are derived is indicated in the "line" column, with a dash indicating an unknown or unnamed line. The "Ph." (phenotype) column of table 2 indicates whether a given line has been reported as resistant (R) to at least one race of SCN or sensitive (S).

The nucleotide bases located at each of 30 positions in each of the haplotype sequences is shown in the columns labeled "Base number of contig 240O17\_region\_G3 of reference line A3244." The base number at the top of each column corresponds to the base number in contig 240O17\_region\_G3 of reference line A3224 (SEQ ID NOs: 2 and 3). The letters G, A, C, and T correspond to the bases guanine, adenine, cytosine, and thymine. Two bases separated by a slash (A/G, C/T, or T/C) indicate uncertainty at the specified position of the haplotype sequence. A "d" followed by a number indicates a deletion of a the length specified. That is, d1 is a one base deletion, d2 is a two base deletion, d14 is a fourteen base deletion, and d19 is a nineteen base deletion. A zero (0) indicates no deletion. A dash indicates that the identity of the base is undetermined.

Examination of table 2 reveals that the amino acid substitutions in the *rhg1* coding region are common to the resistant lines PI467312 (Cha-mo-shi-dou), PI88788 and the southern susceptible lines Essex, Hutchenson, Noir and A1923. As used herein, a "southern" cultivar is any cultivar from maturity groups VI, VII, VIII, IX, or X, and a "northern" cultivar is any cultivar from maturity groups 000, 00, 0, I, II, III, IV, or V. This data is consistent with the mapping experiments of Qui *et al.* (*Theor Appl Genet* 98:356-364 (1999)). Based on analysis of 200 F<sub>2:3</sub> families derived from a cross between Peking and Essex, the authors failed to detect any significant association with SCN resistance to races 1, 2, and 3, and the *rhg1* locus on linkage group G. The authors point out that one of the markers, Bng122, which has been shown to have significant linkage to *rhg1* (Concibido *et al.*, *Crop Sci.* 36:1643-1650 (1996)), is not polymorphic in the

population employed. It is also possible that the susceptible southern lines contain *rhg1* and the susceptible phenotype reflects the polygenic nature of SCN resistance. In a study to uncover QTLs for sudden death syndrome (SDS) in soybean, two SCN resistant alleles originating from the susceptible parent Essex have been described (Hnetkovsky *et al.*, *Crop Sci.* 36:393-400).

Tables 3a, 3b, and 3c, below, show lines that share an *rhg1* haplotype.

**TABLE 3a**

Haplotype 2 Lines		
PI#	Line	Ph.
PI548402	Peking	R
PI404166	Krasnoarmejkaja	R
PI404198	(Sun huan do)	R
PI437654	Er-hej-jan	R
PI438489	(Chiquita)	R
PI507354	Tokai 421	R
PI548655	Forrest	R
PI548988	Pickett	R
PI84751	-	R
PI437654	-	R
PI40792	-	-

**TABLE 3b**

Haplotype 4 Lines		
PI#	Line	Ph.
-	Will	S
PI467312	Cha-mo-shi-dou	R
PI88788	-	R

**TABLE 3c**

Haplotype 6 Lines		
PI#	Line	Ph.
-	Noir	S
-	A1923	S
-	Hutcheson	S

In Tables 3a, 3b, and 3c, Plant Introduction classification number is indicated in the "PI#" column. A dash indicates that no PI number is known or assigned for the line in question. The line from which the sequences are derived is indicated in the "line"



column, with a dash indicating an unknown or unnamed line. The "Ph." column indicates whether a given line has been reported as resistant (R) to at least one race of SCN or sensitive (S), with a dash indicating that the phenotype is unknown.

In a preferred aspect, the source of either an *rhg1* SCN sensitive allele or an *rhg1* SCN resistant allele, or more preferably both, is an elite plant. An "elite line" is any line that has resulted from breeding and selection for superior agronomic performance. Examples of elite lines are lines that are commercially available to farmers or soybean breeders such as HARTZ™ variety H4994, HARTZ™ variety H5218, HARTZ™ variety H5350, HARTZ™ variety H5545, HARTZ™ variety H5050, HARTZ™ variety H5454, HARTZ™ variety H5233, HARTZ™ variety H5488, HARTZ™ variety HLA572, HARTZ™ variety H6200, HARTZ™ variety H6104, HARTZ™ variety H6255, HARTZ™ variety H6586, HARTZ™ variety H6191, HARTZ™ variety H7440, HARTZ™ variety H4452 Roundup Ready™, HARTZ™ variety H4994 Roundup Ready™, HARTZ™ variety H4988 Roundup Ready™, HARTZ™ variety H5000 Roundup Ready™, HARTZ™ variety H5147 Roundup Ready™, HARTZ™ variety H5247 Roundup Ready™, HARTZ™ variety H5350 Roundup Ready™, HARTZ™ variety H5545 Roundup Ready™, HARTZ™ variety H5855 Roundup Ready™, HARTZ™ variety H5088 Roundup Ready™, HARTZ™ variety H5164 Roundup Ready™, HARTZ™ variety H5361 Roundup Ready™, HARTZ™ variety H5566 Roundup Ready™, HARTZ™ variety H5181 Roundup Ready™, HARTZ™ variety H5889 Roundup Ready™, HARTZ™ variety H5999 Roundup Ready™, HARTZ™ variety H6013 Roundup Ready™, HARTZ™ variety H6255 Roundup Ready™, HARTZ™ variety H6454 Roundup Ready™, HARTZ™ variety H6686 Roundup Ready™, HARTZ™ variety H7152 Roundup Ready™, HARTZ™ variety H7550 Roundup Ready™, HARTZ™ variety H8001 Roundup Ready™ (HARTZ SEED, Stuttgart, Arkansas, U.S.A.); A0868, AG0901, A1553, A1900, AG1901, A1923, A2069, AG2101, AG2201, A2247, AG2301, A2304, A2396, AG2401, AG2501, A2506, A2553, AG2701, AG2702, A2704, A2833, A2869, AG2901, AG2902, AG3001, AG3002,

A3204, A3237, A3244, AG3301, AG3302, A3404, A3469, AG3502, A3559, AG3601, AG3701, AG3704, AG3750, A3834, AG3901, A3904, A4045 AG4301, A4341, AG4401, AG4501, AG4601, AG4602, A4604, AG4702, AG4901, A4922, AG5401, A5547, AG5602, A5704, AG5801, AG5901, A5944, A5959, AG6101, QR4459 and QP4544 (Asgrow Seeds, Des Moines, Iowa, U.S.A.); DeKalb variety CX445 (DeKalb, Illinois). An elite plant is any plant from an elite line.

**B) *Rhg4***

The present invention provides a method for the production of a soybean plant having an *Rhg4* SCN resistant allele comprising: (A) crossing a first soybean plant having an *Rhg4* SCN resistant allele with a second soybean plant having an *Rhg4* SCN sensitive allele to produce a segregating population; (B) screening the segregating population for a member having an *Rhg4* SCN resistant allele with a first nucleic acid molecule capable of specifically hybridizing to linkage group A2, wherein the first nucleic acid molecule specifically hybridizes to a second nucleic acid molecule linked to the *Rhg4* SCN resistant allele; and, (C) selecting the member for further crossing and selection.

*Rhg4* is located on linkage group A2 (Matson *et al.*, *Crop Sci.* 5:447 (1965)). SCN resistant alleles of *Rhg4* provide partial resistance to SCN races 1 and 3 (U.S. Patent 5,491,081). Together, *rhg1* and *Rhg4* provide complete or nearly complete resistance to SCN race 3. The dominant gene, *Rhg4*, was found to be closely linked to the seed coat color locus (*i*) (Matson *et al.*, *Crop Sci.* 5:447 (1965)). The *i* locus in Peking was also reported to be linked with a recessive gene for resistance to SCN (Sugiyama *et al.*, *Jpn. J. Breed.* 16:83-86 (1966)). It is possible that *Rhg4* and the recessive gene linked to the *i* locus are one and the same, which would call into question the classification of *Rhg4* as a dominant gene.

Using bioinformatic approaches the *Rhg4* coding region is predicted to contain 2 exons (coding coordinates 111805-113968 and 114684-115204 of SEQ ID NO: 4). *Rhg4* encodes an 894 amino acid polypeptide. *Rhg4* codes for a Xa21-like receptor kinase (SEQ ID NOs: 1099 and 1116-1119) (Song *et al.*, *Science* 270, 1804-1806, (1995)).

*Rhg4* has an extracellular LRR domain (*Rhg4*, SEQ ID NO: 1099, residues 34-44), a transmembrane domain (*Rhg4* SEQ ID NO: 1099, residues 449-471), and STK domain (*Rhg4*, SEQ ID NO: 1099, residues 531-830). In a preferred embodiment, the LRR domain has multiple LRR repeats. In a more preferred embodiment, the LRR domain has 12 LRR repeats.

To identify proteins similar to the *Rhg4* candidate, database searches are performed using the predicted peptide sequences. The *Rhg4* candidate shows similarity to TMK (Y07748)(73.0% similarity and 54.8% identity (CLUSTALW (default parameters))) and TMK1 PRECURSOR (70.6% similarity and 55.1% identity (CLUSTALW (default parameters))), which are rice and *Arabidopsis* receptor kinases, respectively. The predicted LRR extracellular domain reveals similarity to TMK (Y07748)(70.1% similarity and 46.6% identity (CLUSTALW (default parameters))), TMK1 PRECURSOR (g1707642) (65.8% similarity and 48.8% identity (CLUSTALW (default parameters))), and F21J9.1 (g2213607) (65.5% similarity and 45.6% identity (CLUSTALW (default parameters))).

Figure 2 is an alignment of the LRR domain of the *Rhg4* gene. A consensus sequence is shown as the top row. Each row of amino acids represents an LRR domain. The boxed region indicates the putative  $\beta$ -turn/  $\beta$ -sheet structural motif postulated to be involved in ligand binding (Jones *et al.*, *Adv. Bot. Res. Incorpor. Adv. Plant Path.* 24;89-167 (1997)). The hydrophobic leucine residues are thought to project into the core of the protein while the flanking amino acids are thought to be solvent exposed where they may interact with the ligand (Kobe *et al.*, *Nature* 374; 183-186 (1995)). An "x" represents an arbitrary amino acid while an "a" represents a hydrophobic residue (leucine, isoleucine, methionine, valine, or phenylalanine). Amino acid substitutions between resistant and sensitive phenotypes are bordered by a double line. The amino acid substitution within the 35-57 region is a histidine/glutamine substitution, and the amino acid substitution within the 81-104 region is a leucine/phenylalanine substitution.

As used herein, a naturally-occurring *Rhg4* allele is any allele that encodes for a protein having an extracellular LRR domain, a transmembrane domain, and STK domain where the naturally occurring allele is present on linkage group A2 and where certain *Rgh4* alleles, but not all *Rgh4* alleles, are capable of providing or contributing to resistance or partial resistance to a race of SCN. It is understood that such an allele can, using, for example methods disclosed herein, be manipulated so that the nucleic acid molecule encoding the protein is no longer present on linkage group A2. It is also understood that such an allele can, using, for example methods disclosed herein, be manipulated so that the nucleic acid molecule sequence is altered.

As used herein, an *Rhg4* SCN resistant allele is any *Rhg4* allele where that allele alone or in combination with other SCN resistant alleles present in the plant, such as an *rhg1* SCN resistant allele, provides resistance to a race of SCN, and that resistance is due, at least in part, to the genetic contribution of the *Rhg4* allele.

Any soybean plant having an *Rhg4* SCN resistant allele can be used in conjunction with the present invention. Soybeans with known *Rhg4* SCN resistant alleles can be used. Such soybeans include, but are not limited to, PI548402 (Peking), PI437654 (Er-hej-jan), PI438489 (Chiquita), PI507354 (Tokei 421), PI548655 (Forrest), PI548988 (Pickett), PI88788, PI404198 (Sun Huan Do), PI404166 (Krasnoarmejkaja), Hartwig, Manokin, Doles, Dyer, and Custer. In a preferred aspect, the soybean plant having an *Rhg4* SCN resistant allele is an *Rhg4* haplotype 3 allele in a plant having either an *rhg1* haplotype 2 or *rhg1* haplotype 4 allele. Examples of soybeans with an *Rhg4* haplotype 3 allele are PI548402 (Peking), PI88788, PI404198 (Sun huan do), PI438489 (Chiquita), PI437654 (Er-hej-jan), PI404166 (Krasnoarmejkaja), PI548655 (Forrest), PI548988 (Pickett), and PI507354 (Tokei 421). In addition, using the methods or agents of the present invention, soybeans and wild relatives of soybeans such as *Glycine soja* can be screened for the presence of *Rhg4* SCN resistant alleles.

Table 4 below is a table showing single nucleotide polymorphisms (SNPs) for three haplotype sequences of *Rhg4*.

TABLE 4

Identification					Base number of contig 318O13_region_A3					Markers		
Hap	PI number	Line	Ph	Coat	111933	112065	112101	112461	114066	scn279	scn267	scn273
1	-	A2069	R	yellow	T	A	T	A	T	2	2	2
1	-	A2869	R	yellow	T	A	T	A	T	2	2	2
1	-	A3244	S	yellow	T	A	T	A	T	2	2	2
1	PI87631	Kindaizu	R	yellow	T	A	T	A	T	2	2	2
1	PI548389	Minsoy	S	yellow	T	A	T	A	T	2	2	2
1	PI518664	Hutcheson	S	yellow	T	A	T	A	T	2	2	2
1	PI548658	Lee 74	S	yellow	T	A	T	A	T	-	2	2
2	PI540556	Jack	R	yellow	G	A	T	A	T	2	2	1
2	PI360843	Oshimashirome	R	yellow	G	A	T	A	T	-	-	-
2	PI423871	Toyosuzu	R	yellow	G	A	T	A	T	-	-	-
3	PI548402	Peking	R	black	G	C	C	T	G	1	1	1
3	PI88788	-	R	black	G	C	C	T	G	1	1	1
3	PI404198 B	(Sun huan do)	R	black	G	C	C	T	G	1	1	1
3	PI438489 B	(Chiquita)	R	black	G	C	C	T	G	1	1	1
3	PI437654	Er-hej-jan	R	black	G	C	C	T	G	2	1	1
3	PI404166	Krasnoarmeikaja	R	black	G	C	C	T	G	1	1	-
3	PI290136	Noir	S	black	G	C	C	T	G	1	1	1
3	PI548655	Forrest	R	yellow	G	C	C	T	G	1	1	1
3	PI548988	Pickett	R	yellow	G	C	C	T	G	1	1	1
3	PI507354	Takei 421	R	yellow	G	C	C	T	G	1	1	1
N/A	PI467312	Cha-mo-shi-don	R	GnBr	G	C	C	T	-	1	1	1
N/A	PI209332	No.4	R	black	T	A	T	-	-	2	2	2
N/A	PI518672	Will	S	yellow	T	A	T	-	T	2	2	2
N/A	PI548667	Essex	S	yellow	T	A	T	-	T	2	2	2

In Table 4, discrete haplotypes are designated 1 through 3. N/A refers to a haplotype that is not characterized. In Table 4, the Plant Introduction classification number is indicated in the "PI#" column. A dash indicates that no PI number is known or assigned for the line under investigation. The line from which the sequences are derived is indicated in the "line" column, with a dash indicating an unknown or unnamed line. The "Ph." column of Table 4 indicates whether a given line has been reported to be resistant (R) to at least one race of SCN, or sensitive (S). The "coat" column shows the phenotypic coat color of a seed as either yellow, black, green/brown (GnBr), or unknown/unassigned (dash). At the *I* locus, black seeded varieties harbor the *i* allele for black or imperfect black seed coat. In a preferred embodiment, the seed has a yellow coat.

The nucleotide base located at each of 5 positions in each of the haplotype sequences is shown in the columns labeled "Base number of contig 318O13\_region\_A3."

The base number at the top of each column correspond to the base number in the contig 318O13\_region\_A3 of reference line A3244 (SEQ ID NO: 4). The letters G, A, C, and T correspond to the bases guanine, adenine, cytosine, and thymine. A dash indicates that the identity of the base is unknown.

Three different simple sequence repeat (SSR) or microsatellite markers that occur within the sequences, scn279 (SEQ ID NO: 292), scn267 (SEQ ID NO: 282), and scn273 (SEQ ID NO: 294), are listed under "markers." The allele of each marker occurring in a haplotype is indicated by a 1 or a 2, with a dash indicating that the information is not determined.

Any soybean plant having an *Rhg4* SCN sensitive allele can be used in conjunction with the present invention. Such soybeans include A3244, Will, Noir, Lee 74, Essex, Minsoy, A2704, A2833, AG3001, Williams, DK23-51, and Hutcheson. In a preferred aspect, the soybean plant having an *Rhg4* SCN sensitive allele is an *Rhg4* A3244 allele. In addition, using the methods or agents of the present invention, soybeans and wild relative of soybean such as *Glycine soja* can be screened for the presence of *Rhg4* SCN sensitive alleles.

In a preferred aspect, the source of either an *Rhg4* SCN sensitive allele or an *Rhg4* SCN resistant allele, or more preferably both, is an elite plant.

In table 5, below, *rhg1* and *Rhg4* haplotypes for various cultivars are compared.

TABLE 5

Identification				Haplotype	
PI#	Line	Coat	Ph.	<i>rhg4</i>	<i>rhg1</i>
-	A3244	yellow	S	1	1
PI548402	Peking	black	R	3	2
PI404198 B	(Sun huan do)	black	R	3	2
PI438489 B	(Chiquita)	black	R	3	2
PI437654	Er-hej-jan	black	R	3	2
PI404166	Krasnoarmejkaja	black	R	3	2
PI548655	Forrest	yellow	R	3	2
PI548988	Pickett	yellow	R	3	2
PI507354	Tokei 421	yellow	R	3	2
PI88788	-	black	R	3	4

Identification				Haplotype	
PI#	Line	Coat	Ph.	<i>rhg4</i>	<i>rhg1</i>
PI467312	Cha-mo-shi-dou	GnBr	R	N/A	4
-	Noir	black	S	3	6
-	Jack	yellow	R	2	N/A
PI360843	Oshimashirome	yellow	R	2	N/A
PI423871	Toyosuzu	yellow	R	2	3
PI209332	No.4	black	R	N/A	N/A
PI87631	Kindaizu	yellow	R	1	-
-	Minsoy	yellow	S	1	N/A
-	Will	yellow	S	N/A	4
-	Hutcheson	yellow	S	1	6
-	Lee 74	yellow	S	N/A	7
-	Essex	yellow	S	N/A	N/A
-	A2069	yellow	R	1	N/A
-	A2869	yellow	R	1	N/A

In Table 5, haplotypes, as used in Tables 2 through 4, are listed for each line.

N/A refers to a haplotype that is not characterized. The Plant Introduction classification number is indicated in the "PI#" column. A dash indicates that no PI number is known or assigned for the line under investigation. The line from which the sequences are derived is indicated in the "line" column, with a dash indicating an unknown or unnamed line.

The "Ph." column of table 5 indicates whether a given line has been reported to be resistant (R) to at least one race of SCN, or sensitive (S). The "coat" column shows the phenotypic coat color of a seed as either yellow, black, green/brown (GnBr), or unknown/unassigned (dash). At the *I* locus, black seeded varieties harbor the *i* allele for black or imperfect black seed coat. In a preferred embodiment, the seed has a yellow coat.

#### Screening for *rhg1* and *Rhg4* alleles

Any appropriate method can be used to screen for a plant having an *rhg1* SCN resistant allele. Any appropriate method can be used to screen for a plant having an *Rhg4* SCN resistant allele. In a preferred aspect of the present invention, a nucleic acid marker of the present invention can be used (see section entitled "Screening for *rhg1* and *Rhg4* alleles" and subsection (ii) of the section entitled "Agents").

Additional markers, such as SSRs, AFLP markers, RFLP markers, RAPD markers, phenotypic markers, SNPs, isozyme markers, microarray transcription profiles that are genetically linked to or correlated with alleles of a QTL of the present invention can be utilized (Walton, *Seed World* 22-29 (July, 1993); Burow *et al.*, *Molecular Dissection of Complex Traits*, 13-29, Eds. Paterson, CRC Press, New York (1988)). Methods to isolate such markers are known in the art. For example, locus-specific SSRs can be obtained by screening a genomic library for SSRs, sequencing of "positive" clones, designing primers which flank the repeats, and amplifying genomic DNA with these primers. The size of the resulting amplification products can vary by integral numbers of the basic repeat unit. To detect a polymorphism, PCR products can be radiolabeled, separated on denaturing polyacrylamide gels, and detected by autoradiography. Fragments with size differences >4 bp can also be resolved on agarose gels, thus avoiding radioactivity.

Other SSR markers may be utilized. Amplification of simple tandem repeats, mainly of the  $[CA]_n$  type were reported by Litt *et al.*, *Amer. J. Human Genet.* 44:397-401 (1989); Smeets *et al.*, *Human Genet.* 83:245-251 (1989); Tautz, *Nucleic Acids Res.* 17:6463-6472 (1989); Weber *et al.*, *Am. J. Hum. Genet.* 44:388-396 (1989). Weber, *Genomics* 7:524-530 (1990), reported that the level of polymorphism detected by PCR-amplified  $[CA]_n$  type SSRs depends on the number of the "perfect" (*i.e.*, uninterrupted), tandemly repeated motifs. Below a certain threshold (*i.e.*, 12 CA-repeats), the SSRs were reported to be primarily monomorphic. Above this threshold, however, the probability of polymorphism increases with SSR length. Consequently, long, perfect arrays of SSRs are preferred for the generation of markers, *i.e.*, for the design and synthesis of flanking primers.

Suitable primers can be deduced from DNA databases (*e.g.*, Akkaya *et al.*, *Genetics*. 132:1131-1139 (1992)). Alternatively, size-selected genomic libraries (200 to 500 bp) can be constructed by, for example, using the following steps: (1) isolation of genomic DNA; (2) digestion with one or more 4 base-specific restriction enzymes; (3) size-selection of restriction fragments by agarose gel electrophoresis, excision and purification of the desired size fraction; (4) ligation of the DNA into a suitable vector and



transformation into a suitable *E. coli* strain; (5) screening for the presence of SSRs by colony or plaque hybridization with a labeled probe; (6) isolation of positive clones and sequencing of the inserts; and (7) design of suitable primers flanking the SSR.

Establishing libraries with small, size-selected inserts can be advantageous for SSR isolation for two reasons: (1) long SSRs are often unstable in *E. coli*, and (2) positive clones can be sequenced without subcloning. A number of approaches have been reported for the enrichment of SSRs in genomic libraries. Such enrichment procedures are particularly useful if libraries are screened with comparatively rare tri- and tetranucleotide repeat motifs. One such approach has been described by Ostrander *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 89:3419-3423 (1992), who reported the generation of a small-insert phagemid library in an *E. coli* strain deficient in UTPase (d8t) and uracil-N-glycosylase (ung) genes. In the absence of UTPase and uracil-N-glycosylase, dUTP can compete with dTTP for the incorporation into DNA. Single-stranded phagemid DNA isolated from such a library can be primed with [CA]<sub>n</sub> and [TG]<sub>n</sub> primers for second strand synthesis, and the products used to transform a wild-type *E. coli* strain. Since under these conditions there will be selection against single-stranded, uracil-containing DNA molecules, the resulting library will consist of primer-extended, double-stranded products and an about 50-fold enrichment in CA-repeats.

Other reported enrichment strategies rely on hybridization selection of simple sequence repeats prior to cloning (Karagoyozov *et al.*, *Nucleic Acids Res.* 21:3911-3912 (1993); Armour *et al.*, *Hum. Mol. Gen.* 3:599-605 (1994); Kijas *et al.*, *Genome* 38:349-355 (1994); Kandpal *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 91:88-92 (1994); Edwards *et al.*, *Am. J. Hum. Genet.* 49:746-756 (1991)). Hybridization selection, can for example, involve the following steps: (1) genomic DNA is fragmented, either by sonication, or by digestion with a restriction enzyme; (2) genomic DNA fragments are ligated to adapters that allow a "whole genome PCR" at this or a later stage of the procedure; (3) genomic DNA fragments are amplified, denatured and hybridized with single-stranded SSR sequences bound to a nylon membrane; (4) after washing off unbound DNA, hybridizing fragments enriched for SSRs are eluted from the membrane by boiling or alkali treatment, reamplified using adapter-complementary primers, and digested with a restriction enzyme

to remove the adapters; and (5) DNA fragments are ligated into a suitable vector and transformed into a suitable *E. coli* strain. SSRs can be found in up to 50-70% of the clones obtained from these procedures (Armour *et al.*, *Hum. Mol. Gen.* 3:599-605 (1994); Edwards *et al.*, *Am. J. Hum. Genet.* 49:746-756 (1991)).

An alternative hybridization selection strategy was reported by Kijas *et al.*, *Genome* 38:599-605 (1994), which replaced the nylon membrane with biotinylated, SSR-complementary oligonucleotides attached to streptavidin-coated magnetic particles. SSR-containing DNA fragments are selectively bound to the magnetic beads, reamplified, restriction-digested and cloned.

It is further understood that other additional markers on linkage group G or A2 may be utilized (Morgante *et al.*, *Genome* 37:763-769 (1994)). As used herein, reference to the linkage group of G or A2 refers to the linkage group that corresponds to linkage groups U5 and U3, respectively from the genetic map of *Glycine max* (Mansur *et al.*, *Crop Sci.* 36: 1327-1336 (1996), and linkage groups G and A2, respectively, of *Glycine max* x. *Glycine soja* (Shoemaker *et al.*, *Genetics* 144: 329-336 (1996)) that is present in *Glycine soja* (Soybase, an Agricultural Research Service, United States Department of Agriculture (<http://129.186.26.940/> and USDA - Agricultural Research Service: <http://www.ars.usda.gov/>)).

PCR-amplified SSRs can be used, because they are locus-specific, codominant, occur in large numbers and allow the unambiguous identification of alleles. Standard PCR-amplified SSR protocols use radioisotopes and denaturing polyacrylamide gels to detect amplified SSRs. In many situations, however, allele sizes are sufficiently different to be resolved on high percentage agarose gels in combination with ethidium bromide staining (Bell *et al.*, *Genomics* 19:137-144 (1994); Becker *et al.*, *Genome* 38:991-998 (1995); Huttel, Ph.D. Thesis, University of Frankfurt, Germany (1996)). High resolution without applying radioactivity is also provided by nondenaturing polyacrylamide gels in combination with either ethidium bromide (Scrimshaw, *Biotechniques* 13:2189 (1992)) or silver staining (Klinkicht *et al.*, *Molecular Ecology* 1: 133-134 (1992); Neilan *et al.*, *Biotechniques* 17:708-712 (1994)). An alternative of PCR-amplified SSRs typing involves the use of fluorescent primers in combination with a semi-automated DNA

sequencer (Schwengel *et al.*, *Genomics* 22:46-54 (1994)). Fluorescent PCR products can be detected by real-time laser scanning during gel electrophoresis. An advantage of this technology is that different amplification reactions as well as a size marker (each labeled with a different fluorophore) can be combined into one lane during electrophoresis. Multiplex analysis of up to 24 different SSR loci per lane has been reported (Schwengel *et al.*, *Genomics* 22:46-54 (1994)).

The detection of polymorphic sites in a sample of DNA may be facilitated through the use of nucleic acid amplification methods. Such methods specifically increase the concentration of polynucleotides that span the polymorphic site, or include that site and sequences located either distal or proximal to it. Such amplified molecules can be readily detected by gel electrophoresis or other means.

The most preferred method of achieving such amplification employs the polymerase chain reaction ("PCR") (Mullis *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* 51:263-273 (1986); Erlich *et al.*, European Patent Appln. 50,424; European Patent Appln. 84,796, European Patent Application 258,017, European Patent Appln. 237,362; Mullis, European Patent Appln. 201,184; Mullis *et al.*, U.S. Patent No. 4,683,202; Erlich, U.S. Patent No. 4,582,788; and Saiki *et al.*, U.S. Patent No. 4,683,194), using primer pairs that are capable of hybridizing to the proximal sequences that define a polymorphism in its double-stranded form.

In lieu of PCR, alternative methods, such as the "Ligase Chain Reaction" ("LCR") may be used (Barany, *Proc. Natl. Acad. Sci. (U.S.A.)* 88:189-193 (1991)). LCR uses two pairs of oligonucleotide probes to exponentially amplify a specific target. The sequences of each pair of oligonucleotides is selected to permit the pair to hybridize to abutting sequences of the same strand of the target. Such hybridization forms a substrate for a template-dependent ligase. As with PCR, the resulting products thus serve as a template in subsequent cycles and an exponential amplification of the desired sequence is obtained.

LCR can be performed with oligonucleotides having the proximal and distal sequences of the same strand of a polymorphic site. In one embodiment, either oligonucleotide will be designed to include the actual polymorphic site of the

polymorphism. In such an embodiment, the reaction conditions are selected such that the oligonucleotides can be ligated together only if the target molecule either contains or lacks the specific nucleotide that is complementary to the polymorphic site present on the oligonucleotide. Alternatively, the oligonucleotides may be selected such that they do not include the polymorphic site (see, Segev, PCT Application WO 90/01069).

The "Oligonucleotide Ligation Assay" ("OLA") may alternatively be employed (Landegren *et al.*, *Science* 241:1077-1080 (1988)). The OLA protocol uses two oligonucleotides that are designed to be capable of hybridizing to abutting sequences of a single strand of a target. OLA, like LCR, is particularly suited for the detection of point mutations. Unlike LCR, however, OLA results in "linear" rather than exponential amplification of the target sequence.

Nickerson *et al.* have described a nucleic acid detection assay that combines attributes of PCR and OLA (Nickerson *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 87:8923-8927 (1990)). In this method, PCR is used to achieve the exponential amplification of target DNA, which is then detected using OLA. In addition to requiring multiple, and separate, processing steps, one problem associated with such combinations is that they inherit all of the problems associated with PCR and OLA.

Schemes based on ligation of two (or more) oligonucleotides in the presence of a nucleic acid having the sequence of the resulting "di-oligonucleotide," thereby amplifying the di-oligonucleotide, are also known (Wu *et al.*, *Genomics* 4:560-569 (1989)), and may be readily adapted to the purposes of the present invention.

Other known nucleic acid amplification procedures, such as allele-specific oligomers, branched DNA technology, transcription-based amplification systems, or isothermal amplification methods may also be used to amplify and analyze such polymorphisms (Malek *et al.*, U.S. Patent 5,130,238; Davey *et al.*, European Patent Application 329,822; Schuster *et al.*, U.S. Patent 5,169,766; Miller *et al.*, PCT Patent Application WO 89/06700; Kwok, *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 86:1173-1177 (1989); Gingeras *et al.*, PCT Patent Application WO 88/10315; Walker *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 89:392-396 (1992)).

Polymorphisms can also be identified by Single Strand Conformation Polymorphism (SSCP) analysis. SSCP is a method capable of identifying most sequence variations in a single strand of DNA, typically between 150 and 250 nucleotides in length (Elles, *Methods in Molecular Medicine: Molecular Diagnosis of Genetic Diseases*, Humana Press (1996); Orita *et al.*, *Genomics* 5: 874-879 (1989)). Under denaturing conditions a single strand of DNA will adopt a conformation that is uniquely dependent on its sequence conformation. This conformation usually will be different, even if only a single base is changed. Most conformations have been reported to alter the physical configuration or size sufficiently to be detectable by electrophoresis. A number of protocols have been described for SSCP including, but not limited to, Lee *et al.*, *Anal. Biochem.* 205: 289-293 (1992); Suzuki *et al.*, *Anal. Biochem.* 192: 82-84 (1991); Lo *et al.*, *Nucleic Acids Research* 20: 1005-1009 (1992); Sarkar *et al.*, *Genomics* 13:441-443 (1992). It is understood that one or more of the nucleic acids of the present invention can be utilized as markers or probes to detect polymorphisms by SSCP analysis.

Polymorphisms may also be found using random amplified polymorphic DNA (RAPD) (Williams *et al.*, *Nucl. Acids Res.* 18: 6531-6535 (1990)) and cleaveable amplified polymorphic sequences (CAPS) (Lyamichev *et al.*, *Science* 260: 778-783 (1993)). It is understood that one or more of the nucleic acid molecules of the present invention can be utilized as markers or probes to detect polymorphisms by RAPD or CAPS analysis.

The identification of a polymorphism can be determined in a variety of ways. By correlating the presence or absence of it in a plant with the presence or absence of a phenotype, it is possible to predict the phenotype of that plant. If a polymorphism creates or destroys a restriction endonuclease cleavage site, or if it results in the loss or insertion of DNA (*e.g.*, a variable nucleotide tandem repeat (VNTR) polymorphism), it will alter the size or profile of the DNA fragments that are generated by digestion with that restriction endonuclease. As such, individuals that possess a variant sequence can be distinguished from those having the original sequence by restriction fragment analysis. Polymorphisms that can be identified in this manner are termed "restriction fragment length polymorphisms" ("RFLPs"). RFLPs have been widely used in human and plant

genetic analyses (Glassberg, UK Patent Application 2135774; Skolnick *et al.*, *Cytogen. Cell Genet.* 32:58-67 (1982); Botstein *et al.*, *Ann. J. Hum. Genet.* 32:314-331 (1980); Fischer *et al.* (PCT Application WO90/13668); Uhlen, PCT Application WO90/11369)).

A central attribute of "single nucleotide polymorphisms," or "SNPs" is that the site of the polymorphism is at a single nucleotide. SNPs have certain reported advantages over RFLPs and VNTRs. First, SNPs are more stable than other classes of polymorphisms. Their spontaneous mutation rate is approximately  $10^{-9}$  (Kornberg, *DNA Replication*, W.H. Freeman & Co., San Francisco, 1980), approximately 1,000 times less frequent than VNTRs (U.S. Patent 5,679,524). Second, SNPs occur at greater frequency, and with greater uniformity than RFLPs and VNTRs. As SNPs result from sequence variation, new polymorphisms can be identified by sequencing random genomic or cDNA molecules. SNPs can also result from deletions, point mutations and insertions. Any single base alteration, whatever the cause, can be an SNP. The greater frequency of SNPs means that they can be more readily identified than the other classes of polymorphisms.

SNPs and insertion/deletions can be detected by methods, by any of a variety of methods including those disclosed in U.S. Patents 5,210,015; 5,876,930 and 6,030,787 in which an oligonucleotide probe having reporter and quencher molecules is hybridized to a target polynucleotide. The probe is degraded by 5'  $\rightarrow$  3' exonuclease activity of a nucleic acid polymerase. A useful assay is available from AB Biosystems (850 Lincoln Centre Drive, Foster City, CA) as the Taqman® assay.

Specific nucleotide variations such as SNPs and insertion/deletions can also be detected by labeled base extension methods as disclosed in U.S. Patents 6,004,744; 6,013,431; 5,595,890; 5,762,876; and 5,945,283. These methods are based on primer extension and incorporation of detectable nucleoside triphosphates. The primer is designed to anneal to the sequence immediately adjacent to the variable nucleotide which can be detected after incorporation of as few as one labeled nucleoside triphosphate. US Patent 5,468,613 discloses allele specific oligonucleotide hybridizations where single or multiple nucleotide variations in nucleic acid sequence can be detected in nucleic acids by a process in which the sequence containing the nucleotide variation is

amplified, spotted on a membrane and treated with a labeled sequence-specific oligonucleotide probe.

Such methods also include the direct or indirect sequencing of the site, the use of restriction enzymes where the respective alleles of the site create or destroy a restriction site, the use of allele-specific hybridization probes, the use of antibodies that are specific for the proteins encoded by the different alleles of the polymorphism or by other biochemical interpretation. SNPs can be sequenced by a number of methods. Two basic methods may be used for DNA sequencing, the chain termination method of Sanger *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 74: 5463-5467 (1977), and the chemical degradation method of Maxam *et al.*, *Proc. Nat. Acad. Sci. (U.S.A.)* 74: 560-564 (1977). Automation and advances in technology such as the replacement of radioisotopes with fluorescence-based sequencing have reduced the effort required to sequence DNA (Craxton, *Methods*, 2: 20-26 (1991); Ju *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 92: 4347-4351 (1995); Tabor *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 92: 6339-6343 (1995)). Automated sequencers are available from, for example, Pharmacia Biotech, Inc., Piscataway, New Jersey (Pharmacia ALF), LI-COR, Inc., Lincoln, Nebraska (LI-COR 4,000) and Millipore, Bedford, Massachusetts (Millipore BaseStation).

In addition, advances in capillary gel electrophoresis have also reduced the effort required to sequence DNA and such advances provide a rapid high resolution approach for sequencing DNA samples (Swerdlow *et al.*, *Nucleic Acids Res.* 18:1415-1419 (1990); Smith, *Nature* 349:812-813 (1991); Luckey *et al.*, *Methods Enzymol.* 218:154-172 (1993); Lu *et al.*, *J. Chromatog. A.* 680:497-501 (1994); Carson *et al.*, *Anal. Chem.* 65:3219-3226 (1993); Huang *et al.*, *Anal. Chem.* 64:2149-2154 (1992); Kheterpal *et al.*, *Electrophoresis* 17:1852-1859 (1996); Quesada *et al.*, *Electrophoresis* 17:1841-1851 (1996); Baba, *Yakugaku Zasshi* 117:265-281 (1997), Marino, *Appl. Theor. Electrophor.* 5:1-5 (1995)).

The genetic linkage of marker molecules can be established by a gene mapping model such as, without limitation, the flanking marker model reported by Lander *et al.*, *Genetics*, 121:185-199 (1989), and the interval mapping, based on maximum likelihood methods described by Lander *et al.*, *Genetics*, 121:185-199 (1989), and implemented in

the software package MAPMAKER/QTL (Lincoln *et al.*, *Mapping Genes Controlling Quantitative Traits Using MAPMAKER/QTL*, Whitehead Institute for Biomedical Research, Massachusetts, (1990). Additional software includes Qgene, Version 2.23 (1996), Department of Plant Breeding and Biometry, 266 Emerson Hall, Cornell University, Ithaca, NY). Use of Qgene software is a particularly preferred approach.

A maximum likelihood estimate (MLE) for the presence of a marker is calculated, together with an MLE assuming no QTL effect, to avoid false positives. A  $\log_{10}$  of an odds ratio (LOD) is then calculated as:  $\text{LOD} = \log_{10}(\text{MLE for the presence of a QTL} / \text{MLE given no linked QTL})$ .

The LOD score essentially indicates how much more likely the data are to have arisen assuming the presence of a QTL than in its absence. The LOD threshold value for avoiding a false positive with a given confidence, say 95%, depends on the number of markers and the length of the genome. Graphs indicating LOD thresholds are set forth in Lander *et al.*, *Genetics*, 121:185-199 (1989), and further described by Arús *et al.*, *Plant Breeding*, Hayward *et al.* (eds.) Chapman & Hall, London, pp. 314-331 (1993).

Additional models can be used. Many modifications and alternative approaches to interval mapping have been reported, including the use of non-parametric methods (Kruglyak *et al.*, *Genetics*, 139:1421-1428 (1995)). Multiple regression methods or models can be also be used, in which the trait is regressed on a large number of markers (Jansen, *Biometrics in Plant Breed*, van Oijen, Jansen (eds.) Proceedings of the Ninth Meeting of the Eucarpia Section Biometrics in Plant Breeding, The Netherlands, pp. 116-124 (1994); Weber *et al.*, *Advances in Plant Breeding*, Blackwell, Berlin, 16 (1994)). Procedures combining interval mapping with regression analysis, whereby the phenotype is regressed onto a single putative QTL at a given marker interval, and at the same time onto a number of markers that serve as 'cofactors,' have been reported by Jansen *et al.*, *Genetics*, 136:1447-1455 (1994) and Zeng, *Genetics*, 136:1457-1468 (1994). Generally, the use of cofactors reduces the bias and sampling error of the estimated QTL positions (Utz *et al.*, *Biometrics in Plant Breeding*, van Oijen *et al.* (eds.) Proceedings of the Ninth Meeting of the Eucarpia Section Biometrics in Plant Breeding, The Netherlands, pp.195-204 (1994), thereby improving the precision and efficiency of QTL mapping (Zeng,



*Genetics*, 136:1457-1468 (1994)). These models can be extended to multi-environment experiments to analyze genotype-environment interactions (Jansen *et al.*, *Theo. Appl. Genet.* 91:33-37 (1995)).

Selection of an appropriate mapping or segregation populations is important to map construction. The choice of appropriate mapping population depends on the type of marker systems employed (Tanksley *et al.*, *Molecular mapping plant chromosomes. Chromosome structure and function: Impact of new concepts* J.P. Gustafson *et al.* (eds.), Plenum Press, New York, pp. 157-173 (1988)). Consideration must be given to the source of parents (adapted vs. exotic) used in the mapping population. Chromosome pairing and recombination rates can be severely disturbed (suppressed) in wide crosses (adapted x exotic) and generally yield greatly reduced linkage distances. Wide crosses will usually provide segregating populations with a relatively large array of polymorphisms when compared to progeny in a narrow cross (adapted x adapted).

As used herein, the progeny include not only, without limitation, the products of any cross (be it a backcross or otherwise) between two plants, but all progeny whose pedigree traces back to the original cross. Specifically, without limitation, such progeny include plants that have 12.5% or less genetic material derived from one of the two originally crossed plants. As used herein, a second plant is derived from a first plant if the second plant's pedigree includes the first plant.

An  $F_2$  population is the first generation of selfing after the hybrid seed is produced. Usually a single  $F_1$  plant is selfed to generate a population segregating for all the genes in Mendelian (1:2:1) fashion. Maximum genetic information is obtained from a completely classified  $F_2$  population using a codominant marker system (Mather, *Measurement of Linkage in Heredity*: Methuen & Co., (1938)). In the case of dominant markers, progeny tests (*e.g.*,  $F_3$ ,  $BCF_2$ ) are required to identify the heterozygotes, thus making it equivalent to a completely classified  $F_2$  population. However, this procedure is often prohibitive because of the cost and time involved in progeny testing. Progeny testing of  $F_2$  individuals is often used in map construction where phenotypes do not consistently reflect genotype (*e.g.*, disease resistance) or where trait expression is controlled by a QTL. Segregation data from progeny test populations (*e.g.*,  $F_3$  or  $BCF_2$ )

can be used in map construction. Marker-assisted selection can then be applied to cross progeny based on marker-trait map associations ( $F_2$ ,  $F_3$ ), where linkage groups have not been completely disassociated by recombination events (*i.e.*, maximum disequilibrium).

Recombinant inbred lines (RIL) (genetically related lines; usually  $>F_5$ , developed from continuously selfing  $F_2$  lines towards homozygosity) can be used as a mapping population. Information obtained from dominant markers can be maximized by using RIL because all loci are homozygous or nearly so. Under conditions of tight linkage (*i.e.*, about  $<10\%$  recombination), dominant and co-dominant markers evaluated in RIL populations provide more information per individual than either marker type in backcross populations (Reiter *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 89:1477-1481 (1992)). However, as the distance between markers becomes larger (*i.e.*, loci become more independent), the information in RIL populations decreases dramatically when compared to codominant markers.

Backcross populations (*e.g.*, generated from a cross between a successful variety (recurrent parent) and another variety (donor parent) carrying a trait not present in the former) can be utilized as a mapping population. A series of backcrosses to the recurrent parent can be made to recover most of its desirable traits. Thus a population is created consisting of individuals nearly like the recurrent parent but each individual carries varying amounts or mosaic of genomic regions from the donor parent. Backcross populations can be useful for mapping dominant markers if all loci in the recurrent parent are homozygous and the donor and recurrent parent have contrasting polymorphic marker alleles (Reiter *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 89:1477-1481 (1992)). Information obtained from backcross populations using either codominant or dominant markers is less than that obtained from  $F_2$  populations because one, rather than two, recombinant gametes are sampled per plant. Backcross populations, however, are more informative (at low marker saturation) when compared to RILs as the distance between linked loci increases in RIL populations (*i.e.*, about  $.15\%$  recombination). Increased recombination can be beneficial for resolution of tight linkages, but may be undesirable in the construction of maps with low marker saturation.

Near-isogenic lines (NIL) created by many backcrosses to produce an array of individuals that are nearly identical in genetic composition except for the trait or genomic region under interrogation can be used as a mapping population. In mapping with NILs, only a portion of the polymorphic loci are expected to map to a selected region.

Bulk segregant analysis (BSA) is a method developed for the rapid identification of linkage between markers and traits of interest (Michelmore, *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 88:9828-9832 (1991)). In BSA, two bulked DNA samples are drawn from a segregating population originating from a single cross. These bulks contain individuals that are identical for a particular trait (resistant or sensitive to particular disease) or genomic region but arbitrary at unlinked regions (*i.e.*, heterozygous). Regions unlinked to the target region will not differ between the bulked samples of many individuals in BSA.

Plants generated using a method of the present invention can be part of or generated from a breeding program. The choice of breeding method depends on the mode of plant reproduction, the heritability of the trait(s) being improved, and the type of cultivar used commercially (*e.g.*, F<sub>1</sub> hybrid cultivar, pureline cultivar, etc). Selected, non-limiting approaches, for breeding the plants of the present invention are set forth below. A breeding program can be enhanced using marker assisted selection of the progeny of any cross. It is further understood that any commercial and non-commercial cultivars can be utilized in a breeding program. Factors such as, for example, emergence vigor, vegetative vigor, stress tolerance, disease resistance, branching, flowering, seed set, seed size, seed density, standability, and threshability *etc.* will generally dictate the choice.

For highly heritable traits, a choice of superior individual plants evaluated at a single location will be effective, whereas for traits with low heritability, selection should be based on mean values obtained from replicated evaluations of families of related plants. Popular selection methods commonly include pedigree selection, modified pedigree selection, mass selection, and recurrent selection. In a preferred embodiment a backcross or recurrent breeding program is undertaken.

The complexity of inheritance influences choice of the breeding method. Backcross breeding can be used to transfer one or a few favorable genes for a highly

heritable trait into a desirable cultivar. This approach has been used extensively for breeding disease-resistant cultivars. Various recurrent selection techniques are used to improve quantitatively inherited traits controlled by numerous genes. The use of recurrent selection in self-pollinating crops depends on the ease of pollination, the frequency of successful hybrids from each pollination, and the number of hybrid offspring from each successful cross.

Breeding lines can be tested and compared to appropriate standards in environments representative of the commercial target area(s) for two or more generations. The best lines are candidates for new commercial cultivars; those still deficient in traits may be used as parents to produce new populations for further selection.

One method of identifying a superior plant is to observe its performance relative to other experimental plants and to a widely grown standard cultivar. If a single observation is inconclusive, replicated observations can provide a better estimate of its genetic worth. A breeder can select and cross two or more parental lines, followed by repeated selfing and selection, producing many new genetic combinations.

The development of new soybean cultivars requires the development and selection of soybean varieties, the crossing of these varieties and selection of superior hybrid crosses. The hybrid seed can be produced by manual crosses between selected male-fertile parents or by using male sterility systems. Hybrids are selected for certain single gene traits such as pod color, flower color, seed yield, pubescence color or herbicide resistance which indicate that the seed is truly a hybrid. Additional data on parental lines, as well as the phenotype of the hybrid, influence the breeder's decision whether to continue with the specific hybrid cross.

Pedigree breeding and recurrent selection breeding methods can be used to develop cultivars from breeding populations. Breeding programs combine desirable traits from two or more cultivars or various broad-based sources into breeding pools from which cultivars are developed by selfing and selection of desired phenotypes. New cultivars can be evaluated to determine which have commercial potential.

Pedigree breeding is used commonly for the improvement of self-pollinating crops. Two parents who possess favorable, complementary traits are crossed to produce

an  $F_1$ . An  $F_2$  population is produced by selfing one or several  $F_1$ 's. Selection of the best individuals in the best families is performed. Replicated testing of families can begin in the  $F_4$  generation to improve the effectiveness of selection for traits with low heritability. At an advanced stage of inbreeding (*i.e.*,  $F_6$  and  $F_7$ ), the best lines or mixtures of phenotypically similar lines are tested for potential release as new cultivars.

Backcross breeding has been used to transfer genes for a simply inherited, highly heritable trait into a desirable homozygous cultivar or inbred line, which is the recurrent parent. The source of the trait to be transferred is called the donor parent. The resulting plant is expected to have the attributes of the recurrent parent (*e.g.*, cultivar) and the desirable trait transferred from the donor parent. After the initial cross, individuals possessing the phenotype of the donor parent are selected and repeatedly crossed (backcrossed) to the recurrent parent. The resulting parent is expected to have the attributes of the recurrent parent (*e.g.*, cultivar) and the desirable trait transferred from the donor parent.

The single-seed descent procedure in the strict sense refers to planting a segregating population, harvesting a sample of one seed per plant, and using the one-seed sample to plant the next generation. When the population has been advanced from the  $F_2$  to the desired level of inbreeding, the plants from which lines are derived will each trace to different  $F_2$  individuals. The number of plants in a population declines each generation due to failure of some seeds to germinate or some plants to produce at least one seed. As a result, not all of the  $F_2$  plants originally sampled in the population will be represented by a progeny when generation advance is completed.

In a multiple-seed procedure, soybean breeders commonly harvest one or more pods from each plant in a population and thresh them together to form a bulk. Part of the bulk is used to plant the next generation and part is put in reserve. The procedure has been referred to as modified single-seed descent or the pod-bulk technique.

The multiple-seed procedure has been used to save labor at harvest. It is considerably faster to thresh pods with a machine than to remove one seed from each by hand for the single-seed procedure. The multiple-seed procedure also makes it possible to plant the same number of seeds of a population each generation of inbreeding.

Descriptions of other breeding methods that are commonly used for different traits and crops can be found in one of several reference books (*e.g.*, Fehr, *Principles of Cultivar Development* Vol. 1, pp. 2-3 (1987)).

In a preferred aspect of the present invention the source of the *rhg1* SCN resistant allele for use in a breeding program is derived from a plant selected from the group consisting of PI548402 (Peking), PI200499, A2869, Jack, A2069, PI209332 (No:4), PI404166 (Krasnoarmejkaja), PI404198 (Sun huan do), PI437654 (Er-hej-jan), PI438489 (Chiquita), PI507354 (Tokei 421), PI548655 (Forrest), PI548988 (Pickett), PI84751, PI437654, PI40792, Pyramid, Nathan, AG2201, A3469, AG3901, A3904, AG4301, AG4401, AG4501, AG4601, PION9492, PI88788, Dyer, Custer, Manokin, Doles, and SCN resistant progeny thereof (USDA, Soybean Germplasm Collection, University of Illinois, Illinois). In a more preferred aspect, the source of the *rhg1* SCN resistant allele for use in a breeding program is derived from a plant selected from the group consisting of PI548402 (Peking), PI404166 (Krasnoarmejkaja), PI404198 (Sun huan do), PI437654 (Er-hej-jan), PI438489 (Chiquita), PI507354 (Tokei 421), PI548655 (Forrest), PI548988 (Pickett), PI84751, PI437654, PI40792, and SCN resistant progeny thereof.

In a preferred aspect of the present invention the source of the *rhg1* SCN sensitive allele for use in a breeding program is derived from a plant selected from the group consisting of A3244, A2833, AG3001, Williams, Will, A2704, Noir, DK23-51, Lee 74, Essex, Minsoy, A1923, Hutcheson, and SCN sensitive progeny thereof. In a more preferred aspect, the source of the *rhg1* SCN sensitive allele for use in a breeding program is derived from an A3244 plant, and SCN sensitive progeny thereof.

In a preferred aspect of the present invention the source of the *Rhg4* SCN resistant allele for use in a breeding program is derived from a plant selected from the group consisting of PI548402 (Peking), PI437654 (Er-hej-jan), PI438489 (Chiquita), PI507354 (Tokei 421), PI548655 (Forrest), PI548988 (Pickett), PI88788, PI404198 (Sun Huan Do), PI404166 (Krasnoarmejkaja), Hartwig, Manokin, Doles, Dyer, Custer, and SCN resistant progeny thereof. In a more preferred aspect, the source of the *Rhg4* SCN resistant allele for use in a breeding program is derived from a plant selected from the

group consisting of PI548402 (Peking), PI88788, PI404198 (Sun huan do), PI438489 (Chiquita), PI437654 (Er-hej-jan), PI404166 (Krasnoarmejkaja), PI548655 (Forrest), PI548988 (Pickett), PI507354 (Tokei 421), and SCN resistant progeny thereof.

In a preferred aspect of the present invention the source of the *Rhg4* SCN sensitive allele for use in a breeding program is derived from a plant selected from the group consisting of A3244, Will, Noir, Lee 74, Essex, Minsoy, A2704, A2833, AG3001, Williams, DK23-51, and Hutcheson, and SCN sensitive progeny thereof. In a more preferred aspect, the source of the *Rhg4* SCN sensitive allele for use in a breeding program is derived from an A3244 plant, and SCN sensitive progeny thereof.

As used herein linkage of a nucleic acid sequence with another nucleic acid sequence may be genetic or physical. In a preferred embodiment, a nucleic acid marker is genetically linked to either *rhg1* or *Rhg4*, where the marker nucleic acid molecule exhibits a LOD score of greater than 2.0, as judged by interval mapping, for SCN resistance or partial resistance, preferably where the marker nucleic acid molecule exhibits a LOD score of greater than 3.0, as judged by interval mapping, for SCN resistance or partial resistance, more preferably where the marker nucleic acid molecule exhibits a LOD score of greater than 3.5, as judged by interval mapping, for SCN resistance or partial resistance and even more preferably where the marker nucleic acid molecule exhibits a LOD score of about 4.0, as judged by interval mapping, for SCN resistance or partial resistance based on maximum likelihood methods described by Lander *et al.*, *Genetics*, 121:185-199 (1989), and implemented in the software package MAPMAKER/QTL (default parameters)(Lincoln *et al.*, *Mapping Genes Controlling Quantitative Traits Using MAPMAKER/QTL*, Whitehead Institute for Biomedical Research, Massachusetts, (1990)).

In another embodiment the nucleic acid molecule may be physically linked to either *rhg1* or *Rhg4*. In a preferred embodiment, the nucleic acid marker specifically hybridizes to a nucleic acid molecule having a sequence that is present on linkage group G within 500 kb or 100kb, more preferably within 50kb, even more preferably within 25kb of an *rhg1* allele, where the *rhg1* allele is preferably a sensitive allele, and more preferably a sensitive allele from A3244. In a preferred embodiment the nucleic acid

marker is capable of specifically hybridizing to a nucleic acid molecule having a sequence that is present on linkage group A2 within 500kb or 100kb, more preferably within 50kb, even more preferably within 25kb of an *Rhg4* allele, where the *Rgh4* allele is preferably a sensitive allele, and more preferably a sensitive allele from A3244.

The present invention provides a method of investigating an *rhg1* haplotype of a soybean plant comprising: (A) isolating nucleic acid molecules from the soybean plant; (B) determining the nucleic acid sequence of an *rhg1* allele or part thereof; and, (C) comparing the nucleic acid sequence of the *rhg1* allele or part thereof to a reference nucleic acid sequence.

As used herein, the term "investigating" refers to any method capable of detecting a feature, such as a polymorphism or haplotype. Nucleic acid molecules only need to be isolated from a soybean plant to the degree of purity necessary for the task required or to a greater purity if desired. A person of ordinary skill in the art has available techniques to isolate nucleic acid molecules from plants to a sufficient purity, for example without limitation, to sequence the desired region of the nucleic acid molecule or to carry out a marker assay.

The determination of an *rhg1* or *Rhg4* allele or part thereof may be carried out using any technique. Illustration of such techniques include techniques that provide the nucleic acid sequence for an *rhg1* or *rhg4* allele or part thereof include amplification of a desired allele or part thereof (see, for example, the Examples and SEQ ID NOs: 8-53). In a preferred embodiment, the nucleic acid sequence determined is that of an exon of an *rhg1* allele, more preferably exon 1 or exon 3 of an *rhg1* allele, or of an LRR domain. In another preferred embodiment, a single nucleotide is determined. In another preferred embodiment, the nucleic acid sequence determined is that of an LRR domain.

A comparison of a sequence with a reference sequence can be carried out with any appropriate sequence comparison method.

As used herein, a reference sequence is any *rhg1* allele sequence or consensus sequence. A reference sequence may be a nucleic acid sequence or an amino acid sequence. In a preferred embodiment, the reference sequence is any SCN resistant *rhg1* allele sequence. In a further preferred embodiment, the *rhg1* reference sequence is



selected from the group consisting of SEQ ID NOs: 2, 3, 5, 6, 8-23, 28-43, 1097, 1098, and 1100-1115.

The present invention provides a method of investigating an *Rhg4* haplotype of a soybean plant comprising: (A) isolating nucleic acid molecules from the soybean plant; (B) determining the nucleic acid sequence of an *Rhg4* allele or part thereof; and (C) comparing the nucleic acid sequence of the *Rhg4* allele or part thereof to a reference nucleic acid sequence.

As used herein, a reference sequence is any *Rhg4* allele sequence or consensus sequence. A reference sequence may be a nucleic acid sequence or an amino acid sequence. In a preferred embodiment, the reference sequence is any SCN resistant *Rhg4* allele sequence. In a further preferred embodiment, the *Rhg4* reference sequence is selected from the group consisting of SEQ ID NOs: 4, 7, 44-47, 50-53, 1099, and 1116-1119.

The present invention provides a method of introgressing SCN resistance or partial SCN resistance into a soybean plant comprising: performing marker assisted selection of the soybean plant with a nucleic acid marker, wherein the nucleic acid marker specifically hybridizes with a nucleic acid molecule having a first nucleic acid sequence that is physically linked to a second nucleic acid sequence that is located on linkage group G of soybean A3244, wherein the second nucleic acid sequence is within 500 kb of a third nucleic acid sequence which is capable of specifically hybridizing with the nucleic acid sequence of SEQ ID NO: 5, 6, complements thereof, or fragments thereof; and, selecting the soybean plant based on the marker assisted selection.

The present invention provides a method of introgressing SCN resistance or partial SCN resistance into a soybean plant comprising: performing marker assisted selection of the soybean plant with a nucleic acid marker, wherein the nucleic acid marker specifically hybridizes with a nucleic acid molecule having a first nucleic acid sequence that is physically linked to a second nucleic acid sequence that is located on linkage group A2 of soybean A3244, wherein the second nucleic acid sequence is within 500 kb of a third nucleic acid sequence which is capable of specifically hybridizing with the nucleic acid sequence of SEQ ID NO: 7, complements thereof, or fragments thereof; and,

selecting the soybean plant based on the marker assisted selection. Marker assisted introgression of traits into plants has been reported. Marker assisted introgression involves the transfer of a chromosome region defined by one or more markers from one germplasm to a second germplasm. In a preferred embodiment the introgression is carried out by backcrossing with an *rhg1* or *Rhg4* SCN resistant soybean recurrent parent.

In light of the current disclosure, plant introductions and germplasm can be screened with a marker nucleic acid molecule of the present invention to screen for alleles of *rhg1* or *Rhg4* using one or more of techniques disclosed herein or known in the art.

The present invention also provides for parts of the plants produced by a method of the present invention. Plant parts, without limitation, include seed, endosperm, ovule and pollen. In a particularly preferred embodiment of the present invention, the plant part is a seed.

Plants or parts thereof produced by a method of the present invention may be grown in culture and regenerated. Methods for the regeneration of soybean plants from various tissue types and methods for the tissue culture of soybean are known in the art (See, for example, Widholm *et al.*, *In Vitro Selection and Culture-induced Variation in Soybean*, In Soybean: Genetics, Molecular Biology and Biotechnology, Eds. Verma *et al.*, CAB International, Wallingford, Oxon, England (1996)). Regeneration techniques for plants such as soybean can use as the starting material a variety of tissue or cell types. With soybean in particular, regeneration processes have been developed that begin with certain differentiated tissue types such as meristems, Cartha *et al.*, *Can. J. Bot.* 59:1671-1679 (1981), hypocotyl sections, Cameya *et al.*, *Plant Science Letters* 21: 289-294 (1981), and stem node segments, Saka *et al.*, *Plant Science Letters*, 19: 193-201 (1980); Cheng *et al.*, *Plant Science Letters*, 19: 91-99 (1980). Regeneration of whole sexually mature soybean plants from somatic embryos generated from explants of immature soybean embryos has been reported (Ranch *et al.*, *In Vitro Cellular & Developmental Biology* 21: 653-658 (1985). Regeneration of mature soybean plants from tissue culture by organogenesis and embryogenesis has also been reported (Barwale *et al.*, *Planta* 167: 473-481 (1986); Wright *et al.*, *Plant Cell Reports* 5: 150-154 (1986)).

## Agents

One skilled in the art can refer to general reference texts for detailed descriptions of known techniques discussed herein or equivalent techniques. These texts include *Current Protocols in Molecular Biology* Ausubel, *et al.*, eds., John Wiley & Sons, N.Y. (1989), and supplements through September (1998), *Molecular Cloning, A Laboratory Manual*, Sambrook *et al.*, 2<sup>nd</sup> Ed., Cold Spring Harbor Press, Cold Spring Harbor, New York (1989), *Genome Analysis: A Laboratory Manual 1: Analyzing DNA*, Birren *et al.*, Cold Spring Harbor Press, Cold Spring Harbor, New York (1997); *Genome Analysis: A Laboratory Manual 2: Detecting Genes*, Birren *et al.*, Cold Spring Harbor Press, Cold Spring Harbor, New York (1998); *Genome Analysis: A Laboratory Manual 3: Cloning Systems*, Birren *et al.*, Cold Spring Harbor Press, Cold Spring Harbor, New York (1999); *Genome Analysis: A Laboratory Manual 4: Mapping Genomes*, Birren *et al.*, Cold Spring Harbor Press, Cold Spring Harbor, New York (1999); *Plant Molecular Biology: A Laboratory Manual*, Clark, Springer-Verlag, Berlin, (1997), *Methods in Plant Molecular Biology*, Maliga *et al.*, Cold Spring Harbor Press, Cold Spring Harbor, New York (1995). These texts can, of course, also be referred to in making or using an aspect of the invention. It is understood that any of the agents of the invention can be substantially purified and/or be biologically active and/or recombinant.

### (a) Nucleic Acid Molecules

Nucleic acid molecules of the present invention include, without limitation, nucleic acid molecules having a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1-1096 and complements thereof. A subset of the nucleic acid molecules of the present invention includes nucleic acid molecules that encode a protein or fragment thereof. Another subset of the nucleic acid molecules of the present invention are cDNA molecules. Another subset of the nucleic acid molecules of the present invention includes nucleic acid molecules that are marker molecules. A further subset of the nucleic acid molecules of the present invention are those nucleic acid molecules having promoter sequences.

Fragment nucleic acid molecules may comprise significant portion(s) of, or indeed most of, these nucleic acid molecules. In preferred embodiments, the fragments

may comprise smaller polynucleotides, *e.g.*, oligonucleotides having from about 20 to about 250 nucleotide residues and more preferably, about 20 to about 100 nucleotide residues, or about 40 to about 60 nucleotide residues. In another preferred embodiment, fragment molecules may be at least 15 nucleotides, at least 30 nucleotides, at least 50 nucleotides, or at least 100 nucleotides.

The term “substantially purified,” as used herein, refers to a molecule separated from substantially all other molecules normally associated with it in its native state. More preferably a substantially purified molecule is the predominant species present in a preparation. A substantially purified molecule may be greater than 60% free, preferably 75% free, more preferably 90% free, and most preferably 95% free from the other molecules (exclusive of solvent) present in the natural mixture. The term “substantially purified” is not intended to encompass molecules present in their native state.

The agents of the present invention will preferably be “biologically active” with respect to either a structural attribute, such as the capacity of a nucleic acid to hybridize to another nucleic acid molecule, or the ability of a protein to be bound by an antibody (or to compete with another molecule for such binding). Alternatively, such an attribute may be catalytic and thus involve the capacity of the agent to mediate a chemical reaction or response.

The agents of the present invention may also be recombinant. As used herein, the term recombinant describes (a) nucleic acid molecules that are constructed or modified outside of cells and that can replicate or function in a living cell, (b) molecules that result from the transcription, replication or translation of recombinant nucleic acid molecules, or (c) organisms that contain recombinant nucleic acid molecules or are modified using recombinant nucleic acid molecules.

It is understood that the agents of the present invention may be labeled with reagents that facilitate detection of the agent, *e.g.*, fluorescent labels, (Prober *et al.*, *Science* 238:336-340 (1987); Albarella *et al.*, EP 144914), chemical labels, (Sheldon *et al.*, U.S. Patent 4,582,789; Albarella *et al.*, U.S. Patent 4,563,417), and modified bases, (Miyoshi *et al.*, EP 119448) including nucleotides with radioactive elements, *e.g.*,  $^{32}\text{P}$ ,  $^{33}\text{P}$ ,  $^{35}\text{S}$  or  $^{125}\text{I}$ , such as  $^{32}\text{P}$  dCTP.

It is further understood, that the present invention provides recombinant bacterial, animal, fungal and plant cells and viral constructs comprising the agents of the present invention.

Nucleic acid molecules or fragments thereof of the present invention are capable of specifically hybridizing to other nucleic acid molecules under certain circumstances. As used herein, two nucleic acid molecules are said to be capable of specifically hybridizing to one another if the two molecules are capable of forming an anti-parallel, double-stranded nucleic acid structure. A nucleic acid molecule is said to be the "complement" of another nucleic acid molecule if they exhibit "complete complementarity," *i.e.*, each nucleotide in one sequence is complementary to its base pairing partner nucleotide in another sequence. Two molecules are said to be "minimally complementary" if they can hybridize to one another with sufficient stability to permit them to remain annealed to one another under at least conventional "low-stringency" conditions. Similarly, the molecules are said to be "complementary" if they can hybridize to one another with sufficient stability to permit them to remain annealed to one another under conventional "high-stringency" conditions. Nucleic acid molecules which hybridize to other nucleic acid molecules, *e.g.*, at least under low stringency conditions are said to be "hybridizable cognates" of the other nucleic acid molecules. Conventional stringency conditions are described by Sambrook *et al.*, *Molecular Cloning*, A Laboratory Manual, 2nd Ed., Cold Spring Harbor Press, Cold Spring Harbor, New York (1989) and by Haymes *et al.*, *Nucleic Acid Hybridization*, A Practical Approach, IRL Press, Washington, DC (1985). Departures from complete complementarity are therefore permissible, as long as such departures do not completely preclude the capacity of the molecules to form a double-stranded structure. Thus, in order for a nucleic acid molecule to serve as a primer or probe it need only be sufficiently complementary in sequence to be able to form a stable double-stranded structure under the particular solvent and salt concentrations employed.

Appropriate stringency conditions which promote DNA hybridization, for example, 6.0 X sodium chloride/sodium citrate (SSC) at about 45°C, followed by a wash of 2.0 X SSC at 50°C, are known to those skilled in the art or can be found in *Current*

*Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. For example, the salt concentration in the wash step can be selected from a low stringency of about 2.0 X SSC at 50°C to a high stringency of about 0.2 X SSC at 50°C. In addition, the temperature in the wash step can be increased from low stringency conditions at room temperature, about 22°C, to high stringency conditions at about 65°C. Both temperature and salt may be varied, or either the temperature or the salt concentration may be held constant while the other variable is changed.

In a preferred embodiment, a nucleic acid of the present invention will specifically hybridize to one or more of the nucleic acid molecules set forth in SEQ ID NO: 1 through SEQ ID NO: 1096 or complements thereof under moderately stringent conditions, for example at about 2.0 X SSC and about 65°C.

In a particularly preferred embodiment, a nucleic acid of the present invention will include those nucleic acid molecules that specifically hybridize to one or more of the nucleic acid molecules set forth in SEQ ID NO: 1 through SEQ ID NO: 1096 or complements thereof under high stringency conditions such as 0.2 X SSC and about 65°C.

In one aspect of the present invention, the nucleic acid molecules of the present invention comprise one or more of the nucleic acid sequences set forth in SEQ ID NO: 1 through SEQ ID NO: 1096 or complements thereof or fragments of either. In another aspect of the present invention, one or more of the nucleic acid molecules of the present invention share at least 60% sequence identity with one or more of the nucleic acid sequences set forth in SEQ ID NO: 1 through SEQ ID NO: 1096 or complements thereof or fragments of either. In a further aspect of the present invention, one or more of the nucleic acid molecules of the present invention share at least 70% or more, *e.g.*, at least 80%, sequence identity with one or more of the nucleic acid sequences set forth in SEQ ID NO: 1 through SEQ ID NO: 1096 or complements thereof or fragments of either. In a more preferred aspect of the present invention, one or more of the nucleic acid molecules of the present invention share at least 90% or more, *e.g.*, at least 95% and up to 100% sequence identity with one or more of the nucleic acid sequences set forth in SEQ ID NO: 1 through SEQ ID NO: 1096 complements thereof or fragments of either.

As used herein "sequence identity" refers to the extent to which two optimally aligned polynucleotide or peptide sequences are invariant throughout a window of alignment of components, *e.g.*, nucleotides or amino acids. An "identity fraction" for aligned segments of a test sequence and a reference sequence is the number of identical components which are shared by the two aligned sequences divided by the total number of components in reference sequence segment, *i.e.*, the entire reference sequence or a smaller defined part of the reference sequence. "Percent identity" is the identity fraction times 100.

Useful methods for determining sequence identity are disclosed in Guide to Huge Computers, Martin J. Bishop, ed., Academic Press, San Diego, 1994, and Carillo *et al.*, *SIAM J Applied Math* 48:1073 (1988). More particularly, preferred computer programs for determining sequence identity include the Basic Local Alignment Search Tool (BLAST) programs which are publicly available from National Center Biotechnology Information (NCBI) at the National Library of Medicine, National Institute of Health, Bethesda, Md. 20894; see BLAST Manual, Altschul *et al.*, NCBI, NLM, NIH; Altschul *et al.*, *J. Mol. Biol.* 215:403-410 (1990); version 2.0 or higher of BLAST programs allows the introduction of gaps (deletions and insertions) into alignments; BLASTX can be used to determine sequence identity between a polynucleotide sequence query and a protein sequence database; and, BLASTN can be used to determine sequence identity between between sequences.

For purposes of this invention "percent identity" shall be determined using BLASTX version 2.0.14 (default parameters), BLASTN version 2.0.14, or BLASTP 2.0.14.

A particularly preferred group of nucleic acid sequences are those present in the soybean insert of the clones set forth in table 6 below.

TABLE 6

Line	Names of Clones Containing the Specified Gene		
	<i>Rhg4</i>	<i>rhg1/frag 1</i>	<i>rhg1/frag 2</i>
Forrest	Forrest 1	Forrest 7	Forrest13
Peking	Peking 1	Peking 7	Peking 13
Pickett	Pickett 1	Pickett 7	Pickett 13
PI84751	PI 84751.1	PI 84751.7	PI 84751.13

Line	Names of Clones Containing the Specified Gene		
	<i>Rhg4</i>	<i>rhg1/frag 1</i>	<i>rhg1/frag 2</i>
PI87631	PI 87631.1	PI 87631.7	PI 87631.13
PI87631-1	PI 87631-1.1		PI 87631-1.13
PI88788R	PI 88788R.1	PI 88788R.7	PI 88788R.13
PI89772			PI 89772.13
PI90763		PI 90763.7	PI 90763.13
PI200499	PI 200499.1	PI 200499.7	PI 200499.13
PI209332	PI 209332.1		PI 209332.13
PI404166	PI 404166.1	PI 404166.7	PI 404166.13
PI404198A		PI 404198A.7	PI 404198A.13
PI404198B	PI 404198B.1	PI 404198B.7	PI404198B.13
PI437654	PI 437654.1	PI 437654.7	PI 437654.13
PI438489B	PI 438489.1	PI 438489.7	PI 438489B.13
PI467312	PI 467312.1	PI 467312.7	PI 467312.13
PI507354	PI 507354.1	PI 507354.7	PI 507354.13
PI423871	PI 423871.1	PI 423871.7	PI 423871.13
PI407922		PI 407922.7	PI 407922.13
PI360843	PI 360843.1	PI 360843.7	PI 360843.13
A2869	A2869.1	A 2869.7	A2869.13
A2069	A2069.1		A2069.13
Jack	JACK1		JACK13
Will	WILL1	WILL.7	WILL13
Minsoy	Minsoy1	Minsoy.7	MINSOY13
Noir	Noir1	Noir.7	NOIR13
Hutcheson	Hutcheson1	Hutcheson.7	Hutcheson.13
A1923	A1923.1	A1923.7	A1923.13
A2704		A2704.7	A2704.13
Essex	Essex1	Essex.7	ESSEX13
A3244	A3244.1	A 3244.7	A3244.13
Lee74	Lee74.1	Lee74.7	Lee74.13
PI437654		R107C17.7	R107C17.13

Table 5 shows clones comprising *rhg1* and *Rhg4* sequences. The "Lines" column indicates the cultivar from which the sequence in the clone is derived. The *Rhg4*, *rhg1/frag1*, and *rhg1/frag2* columns show the clones derived from the lines that have the *Rhg4*, *rhg1* fragment 1, or *rhg1* fragment 2, respectively. *Rhg4* is amplified with SEQ ID NOs: 48 and 49, which produces a 3.5 kb product. *rhg1* fragment 1 is amplified with SEQ ID NOs: 24 and 25, which produces a 2.9 kb product, and *rhg1* fragment 2 is amplified with SEQ ID NOs: 26 and 27, which produces a 1.75 kb product. All fragments are subcloned into a pCR4-TOPO vector.



(i) **Nucleic Acid Molecules Encoding Proteins or Fragments Thereof**A) *rhg1*

The present invention includes nucleic acid molecules that code for an *rhg1* protein or fragment thereof. Examples of such nucleic acid molecules include those that code for the proteins set forth in SEQ ID NOs: 1097, 1100, 1098, 1101, and 1102-1115. Examples of illustrative fragment molecules include, without limitation, an extracellular LRR domain (*rhg1*, v.1, SEQ ID NO: 1097, residues 164-457; *rhg1*, v.2, SEQ ID NO: 1098, residues 141-434), a transmembrane domain (*rhg1*, v.1, SEQ ID NO: 1097, residues 508-530; *rhg1*, v.2, SEQ ID NO: 1098, residues 33-51 and 485-507), and an STK domain (*rhg1*, v.1, SEQ ID NO: 1097, residues 578-869; *rhg1*, v.2, SEQ ID NO: 1098, residues 555-846). Examples of illustrative nucleic acid molecules include SEQ ID NOs: 5, 6, 8-23, and 28-43.

B) *Rhg4*

The present invention includes nucleic acid molecules that code for an *Rhg4* protein or fragment thereof. Examples of such nucleic acid molecules include those that code for the proteins set forth in SEQ ID NOs: 1099 and 1116-1119. Examples of illustrative fragment molecules include, without limitation, an extracellular LRR domain (SEQ ID NO: 1099, residues 34-44), a transmembrane domain (SEQ ID NO: 1099, residues 449-471), and an STK domain (SEQ ID NO: 1099, residues 531-830). Examples of illustrative nucleic acid molecules include SEQ ID NOs: 7, 44-47, and 50-53.

C) *Rhg1* and *Rhg4*

In another further aspect of the present invention, nucleic acid molecules of the present invention can comprise sequences which differ from those encoding a protein or fragment thereof in SEQ ID NO: 1097 through SEQ ID NO: 1119 due to fact that the different nucleic acid sequence encodes a protein having one or more conservative amino acid changes. It is understood that codons capable of coding for such conservative amino acid substitutions are known in the art.

It is well known in the art that one or more amino acids in a native sequence can be substituted with another amino acid(s), the charge and polarity of which are similar to

that of the native amino acid, *i.e.*, a conservative amino acid substitution. Conserved substitutions for an amino acid within the native polypeptide sequence can be selected from other members of the class to which the naturally occurring amino acid belongs. Amino acids can be divided into the following four groups: (1) acidic amino acids, (2) basic amino acids, (3) neutral polar amino acids, and (4) neutral nonpolar amino acids. Representative amino acids within these various groups include, but are not limited to: (1) acidic (negatively charged) amino acids such as aspartic acid and glutamic acid; (2) basic (positively charged) amino acids such as arginine, histidine, and lysine; (3) neutral polar amino acids such as glycine, serine, threonine, cysteine, cystine, tyrosine, asparagine, and glutamine; and (4) neutral nonpolar (hydrophobic) amino acids such as alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine.

Conservative amino acid changes within the native polypeptides sequence can be made by substituting one amino acid within one of these groups with another amino acid within the same group. Biologically functional equivalents of the proteins or fragments thereof of the present invention can have ten or fewer conservative amino acid changes, more preferably seven or fewer conservative amino acid changes, and most preferably five or fewer conservative amino acid changes. The encoding nucleotide sequence will thus have corresponding base substitutions, permitting it to encode biologically functional equivalent forms of the proteins or fragments of the present invention.

It is understood that certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with structures such as, for example, antigen-binding regions of antibodies or binding sites on substrate molecules. Because it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid sequence substitutions can be made in a protein sequence and, of course, its underlying DNA coding sequence and, nevertheless, obtain a protein with like properties. It is thus contemplated by the inventors that various changes may be made in the peptide sequences of the proteins or fragments of the present invention, or corresponding DNA sequences that encode said peptides, without appreciable loss of their biological utility or activity. It

is understood that codons capable of coding for such amino acid changes are known in the art.

In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydropathic amino acid index in conferring interactive biological function on a protein is generally understood in the art (Kyte *et al.*, *J. Mol. Biol.* 157, 105-132 (1982). It is accepted that the relative hydropathic character of the amino acid contributes to the secondary structure of the resultant protein, which in turn defines the interaction of the protein with other molecules, for example, enzymes, substrates, receptors, DNA, antibodies, antigens, and the like. In making such changes, the substitution of amino acids whose hydropathic indices are within  $\pm 2$  is preferred, those which are within  $\pm 1$  are particularly preferred, and those within  $\pm 0.5$  are even more particularly preferred.

It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U.S. Patent 4,554,101, states that the greatest local average hydrophilicity of a protein, as govern by the hydrophilicity of its adjacent amino acids, correlates with a biological property of the protein. In a further aspect of the present invention, one or more of the nucleic acid molecules of the present invention differ in nucleic acid sequence from those encoding a peptide set forth in SEQ ID NO: 1097 through SEQ ID NO: 1119 or fragment thereof due to the fact that one or more codons encoding an amino acid has been substituted for a codon that encodes a nonessential substitution of the amino acid originally encoded.

Agents of the invention include nucleic acid molecules that encode at least about a contiguous 10 amino acid region of a protein of the present invention, more preferably at least about a contiguous 11 to 14 or larger amino acid region of a protein of the present invention. It is understood that the present invention includes nucleic acid molecules that specifically hybridize or exhibit a particular identity to the nucleic acid molecules described in (i). *See* (a) above.

(ii) **Nucleic Acid Molecule Markers and Collections of Such Molecules**

One aspect of the present invention concerns nucleic acid molecules of the present invention that can act as markers. As used herein, a “marker” is an indicator for the presence of at least one phenotype or polymorphism, such as single nucleotide polymorphisms (SNPs), cleaveable amplified polymorphic sequences (CAPS), amplified fragment length polymorphisms (AFLPs), restriction fragment length polymorphisms (RFLPs), simple sequence repeats (SSRs), or random amplified polymorphic DNA (RAPDs). A “nucleic acid marker” as used herein means a nucleic acid molecule that is capable of being a marker for detecting a polymorphism or phenotype.

In one embodiment of the present invention, the nucleic acid marker specifically hybridizes to a nucleic acid molecule having a nucleic acid sequence selected from the group SEQ NOs: 1-1096 and complements thereof. In a preferred embodiment, the nucleic acid marker is capable of detecting an *rhg1* SNP or INDEL set forth in table 2. In a preferred embodiment, the nucleic acid marker is capable of detecting an *Rgh4* SNP or INDEL set forth in table 4. In another preferred embodiment the nucleic acid marker is a nucleic acid molecule capable of acting as a PCR primer to amplify an *rhg1* or *Rgh4* coding region. Examples of such primers include, without limitation, nucleic acid molecules having a nucleic acid sequence set forth in SEQ ID NO: 401–1096 and complements thereof. Such primers can be used in pairs to amplify a region (amplicons, e.g., without limitation, SEQ ID NOs: 54-400) that can be further investigated using techniques known in the art such as nucleic acid sequencing. Preferred pairs are those with identical “Seq ID” (*see* Description of the Sequence Listing) except for the fact that one “Seq ID” recites forward primer and one recites reverse primer.

In another embodiment of the present invention, the nucleic acid marker specifically hybridizes to a nucleic acid molecule having a sequence that is present on linkage group G within 500 kb or 100kb, more preferably within 50kb, even more preferably within 25kb of an *rhg1* allele, where the *Rgh4* allele is preferably a sensitive allele, and more preferably a sensitive allele from A3244. In a preferred embodiment the nucleic acid marker specifically hybridizes to a nucleic acid molecule having a sequence

that is present on linkage group A2 within 500kb or 100kb, more preferably within 50kb, even more preferably within 25kb of an *Rhg4* allele, where the *Rhg4* allele is preferably a sensitive allele, and more preferably a sensitive allele from A3244.

As used herein, a "collection of nucleic acid molecules" is a population of nucleic acid molecules where at least two, preferably all, of the nucleic acid molecules differ, at least in part, in their nucleic acid sequence. It is understood, that as used herein, an individual species within a collection of nucleic acid molecules may be physically separate or alternatively not physically separate from one or more other species within the collection of nucleic acid molecules. An example of a situation where individual species may be physically separate but considered a collection of nucleic acid molecules is where more than two species are present in a single location such as an array.

As used herein, where a collection of nucleic acid molecules is a marker for a particular attribute, the level, pattern, occurrence and/or absence of the nucleic acid molecules associated with the attribute are not required to be the same between species of the collection. For example, the increase in the level of a species when in combination with the decrease in a second species could be diagnostic for a particular attribute. In a preferred embodiment of the present invention, the level, pattern, occurrence and/or absence of a nucleic acid molecule and/or collection of nucleic acid molecules of the present invention is a marker for SCN resistance.

In one embodiment, the marker is any nucleic acid molecule that specifically hybridizes to any nucleic acid sequence set forth herein. In another embodiment, the marker is a marker capable of distinguishing among the haplotypes of either *rhg1* or *Rhg4*. In yet another embodiment, more than one marker is used to simultaneously distinguish more than one haplotype. In a preferred embodiment, two, three, four, six, eight, twenty five or fifty or more nucleic acid markers are used simultaneously. In another embodiment, one or more markers that are capable of distinguishing among the haplotypes of *rhg1* and one or more markers that are capable of distinguishing among the haplotypes of *Rhg4* are used together.

(iii) **Nucleic acid molecules having promoter sequences and other regulatory sequences**

The present invention includes nucleic acid molecules that are an *rhg1* or *Rhg4* promoter or fragment thereof. Examples of such nucleic acid molecules include those set forth in SEQ ID NO: 2, upstream of coordinate 45163 and SEQ ID NO: 3, upstream of coordinate 46798. As used herein a promoter is a nucleic acid sequence that when joined with a coding region is capable of expressing the protein or fragment thereof so encoded. In a preferred embodiment the promoter sequence corresponds to between 500 nucleotides and 5,000 nucleotides or between 300 nucleotides and 700 nucleotides of the nucleic acid sequence set forth in SEQ ID NO: 2 between coordinates 45163 and 40163, or SEQ ID NO:3 between coordinates 46798 and 41798, or the nucleic acid sequence set forth in SEQ ID NO: 4 between coordinates 111805 and 106805 Preferred partial promoter regions include the TATA box region, e.g. at coordinates 44234 through 44246 of SEQ ID NO: 2 and at coordinates 107826 through 107835 of SEQ ID NO: 4, and CAAT box region, e.g. at coordinates 106243 through 106259 of SEQ ID NO: 4.

Other regulatory sequences include introns or 3' untranslated regions (3'UTRs) associated with *rhg1* and *Rhg4*. In a preferred embodiment, an intron is selected from a nucleic acid comprising a sequence selected from SEQ ID NO: 2 (*rhg1* v.1 at coordinates 45315-45449, 45510-46940, and 48764-48974), SEQ ID NO: 3 (*rhg1* v.2 at coordinates 48764-48974) and SEQ ID NO: 4 (*Rhg4* at coordinates 113969-114683). In another preferred embodiment, a 3'UTR is located within 5,000 nucleotides, more preferable within 1000 nucleotides in the 3' direction of the last coding nucleotide of either *rhg1* or *Rhg4* (SEQ ID NO: 2, *rhg1* v.1, coordinate 49573, SEQ ID NO: 3, *rhg1*, v.2, coordinate 49573, SEQ ID NO: 4, *Rhg4*, coordinate 115204).

It is understood that the present invention includes nucleic acid molecules that specifically hybridize or exhibit a particular identity to the nucleic acid molecules described in (iii). See (a) above.

**(b) Protein and Peptide Molecules**

A class of agents comprises one or more of the protein or peptide molecules encoded by SEQ ID NO: 1097 through SEQ ID NO:1119 or one or more of the protein or fragment thereof or peptide molecules encoded by other nucleic acid agents of the present invention. As used herein, the term "protein molecule" and "peptide molecule" mean any protein or protein fragment or peptide or polypeptide molecule that comprises ten or more amino acids, preferably at least 11 or 12 or more, more preferably at least 13 or 14 amino acids. It is well known in the art that proteins may undergo modification, including post-translational modifications, such as, but not limited to, disulfide bond formation, glycosylation, phosphorylation, or oligomerization. Thus, as used herein, the terms "protein molecule" and "peptide molecule" include molecules that are modified by any biological or non-biological process. The terms "amino acid" and "amino acids" refer to all naturally occurring L-amino acids. This definition is meant to include norleucine, ornithine, homocysteine, and homoserine.

One or more of the protein or peptide molecules may be produced via chemical synthesis, or more preferably, by expression in a suitable bacterial or eukaryotic host. Suitable methods for expression are described by Sambrook, *et al.*, (In: *Molecular Cloning, A Laboratory Manual, 2nd Edition, Cold Spring Harbor Press, Cold Spring Harbor, New York (1989)*, or similar texts.

Another class of agents comprise protein or peptide molecules encoded by SEQ ID NO: 1097 through SEQ ID NO: 1119 or complements thereof or, fragments or fusions thereof in which non-essential, or not relevant, amino acid residues have been added, replaced, or deleted. An example of such a homolog is a protein homolog of each soybean species, including but not limited to alfalfa, barley, *Brassica*, broccoli, cabbage, citrus, garlic, oat, oilseed rape, onion, canola, flax, pea, peanut, pepper, potato, rye, soybean, strawberry, sugarcane, sugarbeet, soybean, maize, rice, cotton, sorghum, *Arabidopsis*, wheat, pine, fir, eucalyptus, apple, lettuce, peas, lentils, grape, banana, tea, turf grasses, *etc.* Particularly preferred non-soybean plants to utilize for the isolation of homologs would include alfalfa, barley, oat, oilseed rape, canola, ornamentals, sugarcane,

sugarbeet, soybean, maize, rice, cotton, sorghum, *Arabidopsis*, wheat, potato, and turf grasses. Such a homolog can be obtained by any of a variety of methods. Most preferably, as indicated above, one or more of the disclosed sequences (SEQ ID NO: 1 through SEQ ID NO: 1096 or complements thereof) will be used to define a pair of primers that may be used to isolate the protein homolog-encoding nucleic acid molecules from any desired species. Such molecules can be expressed to yield protein homologs by recombinant means.

**(c) Plant Constructs and Plant Transformants**

One or more of the nucleic acid molecules of the invention may be used in plant transformation or transfection. Exogenous genetic material may be transferred into a plant cell and the plant cell regenerated into a whole, fertile or sterile plant. Exogenous genetic material is any genetic material, whether naturally occurring or otherwise, from any source that is capable of being inserted into any organism. In a preferred embodiment the exogenous genetic material includes a nucleic acid molecule of the present invention, preferably a nucleic acid molecule having at least 20 nucleotides of a sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 1096 and complements thereof. In a preferred embodiment, the nucleic acid molecule codes for a protein or fragment thereof described in Section (i). In another preferred embodiment, the nucleic acid molecule is a promoter or fragment thereof described in Section (iii).

Such genetic material may be transferred into either monocotyledons and dicotyledons including, but not limited to tomato, eggplant, maize, soybean, *Arabidopsis*, phaseolus, peanut, alfalfa, wheat, rice, oat, sorghum, rye, tritordeum, millet, fescue, perennial ryegrass, sugarcane, cranberry, papaya, banana, muskmelon, apple, cucumber, dendrobium, gladiolus, chrysanthemum, liliacea, cotton, eucalyptus, sunflower, canola, turfgrass, sugarbeet, coffee and dioscorea (Christou, In: *Particle Bombardment for Genetic Engineering of Plants*, Biotechnology Intelligence Unit. Academic Press, San Diego, California (1996).



In a preferred embodiment, the genetic material is transferred to a soybean.

Preferred soybeans to transfer an *rhg1* SCN resistance allele are selected from the group consisting of PI548402 (Peking), PI200499, A2869, Jack, A2069, PI209332 (No:4), PI404166 (Krasnoarmejkaja), PI404198 (Sun huan do), PI437654 (Er-hej-jan), PI438489 (Chiquita), PI507354 (Tokei 421), PI548655 (Forrest), PI548988 (Pickett), PI84751, PI437654, PI40792, Pyramid, Nathan, AG2201, A3469, AG3901, A3904, AG4301, AG4401, AG4501, AG4601, PION9492, PI88788, Dyer, Custer, Manokin, and Doles.

Preferred soybeans to transfer an *Rhg4* SCN resistance allele are selected from the group consisting of PI548402 (Peking), PI437654 (Er-hej-jan), PI438489 (Chiquita), PI507354 (Tokei 421), PI548655 (Forrest), PI548988 (Pickett), PI88788, PI404198 (Sun Huan Do), PI404166 (Krasnoarmejkaja), Hartwig, Manokin, Doles, Dyer, and Custer.

Transfer of a nucleic acid that encodes for a protein can result in overexpression of that protein in a transformed cell or transgenic plant. One or more of the proteins or fragments thereof encoded by nucleic acid molecules of the invention may be overexpressed in a transformed cell or transformed plant. Such overexpression may be the result of transient or stable transfer of the exogenous genetic material. Such overexpression can also result in SCN resistance to one or more races of SCN.

Exogenous genetic material may be transferred into a host cell by the use of a DNA vector or construct designed for such a purpose. Design of such a vector is generally within the skill of the art (*See, Plant Molecular Biology: A Laboratory Manual*, Clark (ed.), Springer, New York (1997).

A construct or vector may include a plant promoter to express the protein or protein fragment of choice. A number of promoters, which are active in plant cells, have been described in the literature. These include the nopaline synthase (NOS) promoter (Ebert *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 84:5745-5749 (1987), the octopine synthase (OCS) promoter (which are carried on tumor-inducing plasmids of *Agrobacterium tumefaciens*), the caulimovirus promoters such as the cauliflower mosaic virus (CaMV) 19S promoter (Lawton *et al.*, *Plant Mol. Biol.* 9:315-324 (1987), and the CaMV 35S

promoter (Odell *et al.*, *Nature* 313:810-812 (1985), the figwort mosaic virus 35S-promoter, the light-inducible promoter from the small subunit of ribulose-1,5-bisphosphate carboxylase (ssRUBISCO), the Adh promoter (Walker *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 84:6624-6628 (1987), the sucrose synthase promoter (Yang *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 87:4144-4148 (1990), the R gene complex promoter (Chandler *et al.*, *The Plant Cell* 1:1175-1183 (1989), and the chlorophyll a/b binding protein gene promoter, *etc.* These promoters have been used to create DNA constructs that have been expressed in plants; *see, e.g.*, PCT publication WO 84/02913. The CaMV 35S promoters are preferred for use in plants. Promoters known or found to cause transcription of DNA in plant cells can be used in the invention.

For the purpose of expression in source tissues of the plant, such as the leaf, seed, root or stem, it is preferred that the promoters utilized have relatively high expression in these specific tissues. Tissue-specific expression of a protein of the present invention is a particularly preferred embodiment. For this purpose, one may choose from a number of promoters for genes with tissue- or cell-specific or -enhanced expression. Examples of such promoters reported in the literature include the chloroplast glutamine synthetase GS2 promoter from pea (Edwards *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 87:3459-3463 (1990), the chloroplast fructose-1,6-bisphosphatase (FBPase) promoter from wheat (Lloyd *et al.*, *Mol. Gen. Genet.* 225:209-216 (1991), the nuclear photosynthetic ST-LS1 promoter from potato (Stockhaus *et al.*, *EMBO J.* 8:2445-2451 (1989), the STK (PAL) promoter and the glucoamylase (CHS) promoter from *Arabidopsis thaliana*. Also reported to be active in photosynthetically active tissues are the ribulose-1,5-bisphosphate carboxylase (RbcS) promoter from eastern larch (*Larix laricina*), the promoter for the *cab* gene, *cab6*, from pine (Yamamoto *et al.*, *Plant Cell Physiol.* 35:773-778 (1994), the promoter for the Cab-1 gene from wheat (Fejes *et al.*, *Plant Mol. Biol.* 15:921-932 (1990), the promoter for the CAB-1 gene from spinach (Lubberstedt *et al.*, *Plant Physiol.* 104:997-1006 (1994), the promoter for the *cab1R* gene from rice (Luan *et al.*, *Plant Cell.* 4:971-981 (1992), the pyruvate, orthophosphate dikinase (PPDK) promoter from maize (Matsuoka *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 90: 9586-9590 (1993), the promoter for the tobacco Lhcb1\*2

gene (Cerdan *et al.*, *Plant Mol. Biol.* 33:245-255 (1997)), the *Arabidopsis thaliana* SUC2 sucrose-H<sup>+</sup> symporter promoter (Truernit *et al.*, *Planta*. 196:564-570 (1995)), and the promoter for the thylakoid membrane proteins from spinach (psaD, psaF, psaE, PC, FNR, atpC, atpD, cab, rbcS). Other promoters for the chlorophyll a/b-binding proteins may also be utilized in the invention, such as the promoters for LhcB gene and PsbP gene from white mustard (*Sinapis alba*; Kretsch *et al.*, *Plant Mol. Biol.* 28:219-229 (1995)).

For the purpose of expression in sink tissues of the plant, such as the tuber of the potato plant, the fruit of tomato, or the seed of maize, wheat, rice and barley, it is preferred that the promoters utilized in the invention have relatively high expression in these specific tissues. A number of promoters for genes with tuber-specific or -enhanced expression are known, including the class I patatin promoter (Bevan *et al.*, *EMBO J.* 8:1899-1906 (1986); Jefferson *et al.*, *Plant Mol. Biol.* 14:995-1006 (1990)), the promoter for the potato tuber ADPGPP genes, both the large and small subunits, the sucrose synthase promoter (Salanoubat *et al.*, *Gene* 60:47-56 (1987), Salanoubat *et al.*, *Gene* 84:181-185 (1989)), the promoter for the major tuber proteins including the 22 kd protein complexes and proteinase inhibitors (Hannapel, *Plant Physiol.* 101:703-704 (1993)), the promoter for the granule bound starch synthase gene (GBSS) (Visser *et al.*, *Plant Mol. Biol.* 17:691-699 (1991)), and other class I and II patatins promoters (Koster-Topfer *et al.*, *Mol Gen Genet.* 219:390-396 (1989); Mignery *et al.*, *Gene.* 62:27-44 (1988)).

Other promoters can also be used to express a protein or fragment thereof in specific tissues, such as seeds or fruits. The promoter for  $\beta$ -conglycinin (Chen *et al.*, *Dev. Genet.* 10: 112-122 (1989)) or other seed-specific promoters such as the napin and phaseolin promoters, can be used. The zeins are a group of storage proteins found in maize endosperm. Genomic clones for zein genes have been isolated (Pedersen *et al.*, *Cell* 29:1015-1026 (1982)) and the promoters from these clones, including the 15 kD, 16 kD, 19 kD, 22 kD, 27 kD and genes, could also be used. Other promoters known to function, for example, in maize include the promoters for the following genes: *waxy*, *Brittle*, *Shrunken 2*, Branching enzymes I and II, starch synthases, debranching enzymes, oleosins, glutelins and sucrose synthases. A particularly preferred promoter for maize

endosperm expression is the promoter for the glutelin gene from rice, more particularly the Osgt-1 promoter (Zheng *et al.*, *Mol. Cell Biol.* 13:5829-5842 (1993)). Examples of promoters suitable for expression in wheat include those promoters for the ADPglucose pyrosynthase (ADPGPP) subunits, the granule bound and other starch synthase, the branching and debranching enzymes, the embryogenesis-abundant proteins, the gliadins and the glutenins. Examples of such promoters in rice include those promoters for the ADPGPP subunits, the granule bound and other starch synthase, the branching enzymes, the debranching enzymes, sucrose synthases and the glutelins. A particularly preferred promoter is the promoter for rice glutelin, Osgt-1. Examples of such promoters for barley include those for the ADPGPP subunits, the granule bound and other starch synthase, the branching enzymes, the debranching enzymes, sucrose synthases, the hordeins, the embryo globulins and the aleurone specific proteins.

Root specific promoters may also be used. An example of such a promoter is the promoter for the acid chitinase gene (Samac *et al.*, *Plant Mol. Biol.* 25:587-596 (1994)). Expression in root tissue could also be accomplished by utilizing the root specific subdomains of the CaMV35S promoter that have been identified (Lam *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 86:7890-7894 (1989)). Other root cell specific promoters include those reported by Conkling *et al.* (Conkling *et al.*, *Plant Physiol.* 93:1203-1211 (1990)).

Additional promoters that may be utilized are described, for example, in U.S. Patent Nos. 5,378,619; 5,391,725; 5,428,147; 5,447,858; 5,608,144; 5,608,144; 5,614,399; 5,633,441; 5,633,435; and 4,633,436. In addition, a tissue specific enhancer may be used (Fromm *et al.*, *The Plant Cell* 1:977-984 (1989)).

Preferred promoters are those set forth in Section (a)(iii) of Agents.

Constructs or vectors may also include, with the coding region of interest, a nucleic acid sequence that acts, in whole or in part, to terminate transcription of that region. A number of such sequences have been isolated, including the Tr7 3' sequence and the NOS 3' sequence (Ingelbrecht *et al.*, *The Plant Cell* 1:671-680 (1989); Bevan *et al.*, *Nucleic Acids Res.* 11:369-385 (1983)).

A vector or construct may also include regulatory elements. Examples of such include the Adh intron 1 (Callis *et al.*, *Genes and Develop.* 1:1183-1200 (1987)), the sucrose synthase intron (Vasil *et al.*, *Plant Physiol.* 91:1575-1579 (1989)) and the TMV omega element (Gallie *et al.*, *The Plant Cell* 1:301-311 (1989)). These and other regulatory elements may be included when appropriate.

A vector or construct may also include a selectable marker. Selectable markers may also be used to select for plants or plant cells that contain the exogenous genetic material. Examples of such include, but are not limited to: a neomycin phosphotransferase gene (U.S. Patent 5,034,322), which codes for kanamycin resistance and can be selected for using kanamycin, G418, *etc.*; a bar gene which codes for bialaphos resistance; genes which encode glyphosate resistance (U.S. Patents 4,940,835; 5,188,642; 4,971,908; 5,627,061); a nitrilase gene which confers resistance to bromoxynil (Stalker *et al.*, *J. Biol. Chem.* 263:6310-6314 (1988)); a mutant acetolactate synthase gene (ALS) which confers imidazolinone or sulphonylurea resistance (European Patent Application 154,204 (Sept. 11, 1985)); and a methotrexate resistant DHFR gene (Thillet *et al.*, *J. Biol. Chem.* 263:12500-12508 (1988)).

A vector or construct may also include DNA sequence which encodes a transit peptide. Incorporation of a suitable chloroplast transit peptide may also be employed (European Patent Application Publication Number 0218571). Translational enhancers may also be incorporated as part of the vector DNA. DNA constructs could contain one or more 5' non-translated leader sequences which may serve to enhance expression of the gene products from the resulting mRNA transcripts. Such sequences may be derived from the promoter selected to express the gene or can be specifically modified to increase translation of the mRNA. Such regions may also be obtained from viral RNAs, from suitable eukaryotic genes, or from a synthetic gene sequence. For a review of optimizing expression of transgenes, see Koziel *et al.*, *Plant Mol. Biol.* 32:393-405 (1996).

A vector or construct may also include a screenable marker. Screenable markers may be used to monitor expression. Exemplary screenable markers include: a  $\beta$ -glucuronidase or uidA gene (GUS) which encodes an enzyme for which various

chromogenic substrates are known (Jefferson, *Plant Mol. Biol. Rep.* 5:387-405 (1987); Jefferson *et al.*, *EMBO J.* 6:3901-3907 (1987)); an R-locus gene, which encodes a product that regulates the production of anthocyanin pigments (red color) in plant tissues (Dellaporta *et al.*, *Stadler Symposium 11*:263-282 (1988)); a  $\beta$ -lactamase gene (Sutcliffe *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 75:3737-3741 (1978)), a gene which encodes an enzyme for which various chromogenic substrates are known (*e.g.*, PADAC, a chromogenic cephalosporin); a luciferase gene (Ow *et al.*, *Science* 234:856-859 (1986)); a xylE gene (Zukowsky *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 80:1101-1105 (1983)) which encodes a catechol dioxygenase that can convert chromogenic catechols; an  $\alpha$ -amylase gene (Ikatsu *et al.*, *Bio/Technol.* 8:241-242 (1990)); a tyrosinase gene (Katz *et al.*, *J. Gen. Microbiol.* 129:2703-2714 (1983)) which encodes an enzyme capable of oxidizing tyrosine to DOPA and dopaquinone which in turn condenses to melanin; an  $\alpha$ -galactosidase, which will turn a chromogenic  $\alpha$ -galactose substrate.

Included within the terms "selectable or screenable marker genes" are also genes which encode a secretable marker whose secretion can be detected as a means of identifying or selecting for transformed cells. Examples include markers which encode a secretable antigen that can be identified by antibody interaction, or even secretable enzymes which can be detected catalytically. Secretable proteins fall into a number of classes, including small, diffusible proteins which are detectable, (*e.g.*, by ELISA), small active enzymes which are detectable in extracellular solution (*e.g.*,  $\alpha$ -amylase,  $\beta$ -lactamase, phosphinothricin transferase), or proteins which are inserted or trapped in the cell wall (such as proteins which include a leader sequence such as that found in the expression unit of extension or tobacco PR-S). Other possible selectable and/or screenable marker genes will be apparent to those of skill in the art.

There are many methods for introducing transforming nucleic acid molecules into plant cells. Suitable methods are believed to include virtually any method by which nucleic acid molecules may be introduced into a cell, such as by *Agrobacterium* infection or direct delivery of nucleic acid molecules such as, for example, by PEG-mediated transformation, by electroporation or by acceleration of DNA coated particles, etc

(Potrykus, *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 42:205-225 (1991); Vasil, *Plant Mol. Biol.* 25:925-937 (1994)). For example, electroporation has been used to transform maize protoplasts (Fromm *et al.*, *Nature* 312:791-793 (1986)).

Other vector systems suitable for introducing transforming DNA into a host plant cell include but are not limited to binary artificial chromosome (BIBAC) vectors (Hamilton *et al.*, *Gene* 200:107-116 (1997)); and transfection with RNA viral vectors (Della-Cioppa *et al.*, *Ann. N.Y. Acad. Sci.* (1996), 792 (Engineering Plants for Commercial Products and Applications), 57-61). Additional vector systems also include plant selectable YAC vectors such as those described in Mullen *et al.*, *Molecular Breeding* 4:449-457 (1988)).

Technology for introduction of DNA into cells is well known to those of skill in the art. Four general methods for delivering a gene into cells have been described: (1) chemical methods (Graham *et al.*, *Virology* 54:536-539 (1973)); (2) physical methods such as microinjection (Capecchi, *Cell* 22:479-488 (1980)), electroporation (Wong *et al.*, *Biochem. Biophys. Res. Commun.* 107:584-587 (1982); Fromm *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 82:5824-5828 (1985); U.S. Patent No. 5,384,253); and the gene gun (Johnston *et al.*, *Methods Cell Biol.* 43:353-365 (1994)); (3) viral vectors (Clapp, *Clin. Perinatol.* 20:155-168 (1993); Lu *et al.*, *J. Exp. Med.* 178:2089-2096 (1993); Eglitis *et al.*, *Biotechniques* 6:608-614 (1988)); and (4) receptor-mediated mechanisms (Curiel *et al.*, *Hum. Gen. Ther.* 3:147-154 (1992); Wagner *et al.*, *Proc. Natl. Acad. Sci. (USA)* 89:6099-6103 (1992)).

Acceleration methods that may be used include, for example, microprojectile bombardment and the like. One example of a method for delivering transforming nucleic acid molecules to plant cells is microprojectile bombardment. This method has been reviewed by Yang *et al.* (eds.), *Particle Bombardment Technology for Gene Transfer*, Oxford Press, Oxford, England (1994)). Non-biological particles (microprojectiles) that may be coated with nucleic acids and delivered into cells by a propelling force. Exemplary particles include those comprised of tungsten, gold, platinum and the like.

A particular advantage of microprojectile bombardment, in addition to it being an effective means of reproducibly transforming monocots, is that neither the isolation of protoplasts (Cristou *et al.*, *Plant Physiol.* 87:671-674 (1988)) nor the susceptibility of *Agrobacterium* infection are required. An illustrative embodiment of a method for delivering DNA into maize cells by acceleration is a biolistics  $\alpha$ -particle delivery system, which can be used to propel particles coated with DNA through a screen, such as a stainless steel or Nytex screen, onto a filter surface covered with corn cells cultured in suspension. Gordon-Kamm *et al.*, describes the basic procedure for coating tungsten particles with DNA (Gordon-Kamm *et al.*, *Plant Cell* 2:603-618 (1990)). The screen disperses the tungsten nucleic acid particles so that they are not delivered to the recipient cells in large aggregates. A particle delivery system suitable for use with the invention is the helium acceleration PDS-1000/He gun is available from Bio-Rad Laboratories (Bio-Rad, Hercules, California)(Sanford *et al.*, *Technique* 3:3-16 (1991)).

For the bombardment, cells in suspension may be concentrated on filters. Filters containing the cells to be bombarded are positioned at an appropriate distance below the microprojectile stopping plate. If desired, one or more screens are also positioned between the gun and the cells to be bombarded.

Alternatively, immature embryos or other target cells may be arranged on solid culture medium. The cells to be bombarded are positioned at an appropriate distance below the microprojectile stopping plate. If desired, one or more screens are also positioned between the acceleration device and the cells to be bombarded. Through the use of techniques set forth herein one may obtain up to 1000 or more foci of cells transiently expressing a screenable or selectable marker gene. The number of cells in a focus which express the exogenous gene product 48 hours post-bombardment often range from one to ten and average one to three.

In bombardment transformation, one may optimize the pre-bombardment culturing conditions and the bombardment parameters to yield the maximum numbers of stable transformants. Both the physical and biological parameters for bombardment are important in this technology. Physical factors are those that involve manipulating the



DNA/microprojectile precipitate or those that affect the flight and velocity of either the macro- or microprojectiles. Biological factors include all steps involved in manipulation of cells before and immediately after bombardment, the osmotic adjustment of target cells to help alleviate the trauma associated with bombardment and also the nature of the transforming DNA, such as linearized DNA or intact supercoiled plasmids. It is believed that pre-bombardment manipulations are especially important for successful transformation of immature embryos.

In another alternative embodiment, plastids can be stably transformed. Methods disclosed for plastid transformation in higher plants include the particle gun delivery of DNA containing a selectable marker and targeting of the DNA to the plastid genome through homologous recombination (Svab *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 87:8526-8530 (1990); Svab *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 90:913-917 (1993); Staub *et al.*, *EMBO J.* 12:601-606 (1993); U.S. Patents 5,451,513 and 5,545,818).

Accordingly, it is contemplated that one may wish to adjust various aspects of the bombardment parameters in small-scale studies to fully optimize the conditions. One may particularly wish to adjust physical parameters such as gap distance, flight distance, tissue distance and helium pressure. One may also minimize the trauma reduction factors by modifying conditions which influence the physiological state of the recipient cells and which may therefore influence transformation and integration efficiencies. For example, the osmotic state, tissue hydration and the subculture stage or cell cycle of the recipient cells may be adjusted for optimum transformation. The execution of other routine adjustments will be known to those of skill in the art in light of the present disclosure.

*Agrobacterium*-mediated transfer is a widely applicable system for introducing genes into plant cells because the DNA can be introduced into whole plant tissues, thereby bypassing the need for regeneration of an intact plant from a protoplast. The use of *Agrobacterium*-mediated plant integrating vectors to introduce DNA into plant cells is well known in the art. See, for example the methods described by Fraley *et al.*, *Bio/Technology* 3:629-635 (1985) and Rogers *et al.*, *Methods Enzymol.* 153:253-277 (1987). Further, the integration of the T-DNA is a relatively precise process resulting in

few rearrangements. The region of DNA to be transferred is defined by the border sequences and intervening DNA is usually inserted into the plant genome as described (Spielmann *et al.*, *Mol. Gen. Genet.* 205:34 (1986)).

Modern *Agrobacterium* transformation vectors are capable of replication in *E. coli* as well as *Agrobacterium*, allowing for convenient manipulations as described (Klee *et al.*, In: *Plant DNA Infectious Agents*, Hohn *et al.* (eds.), Springer-Verlag, New York, pp. 179-203 (1985)). Moreover, technological advances in vectors for *Agrobacterium*-mediated gene transfer have improved the arrangement of genes and restriction sites in the vectors to facilitate construction of vectors capable of expressing various polypeptide coding genes. The vectors described have convenient multi-linker regions flanked by a promoter and a polyadenylation site for direct expression of inserted polypeptide coding genes and are suitable for present purposes (Rogers *et al.*, *Methods Enzymol.* 153:253-277 (1987)). In addition, *Agrobacterium* containing both armed and disarmed Ti genes can be used for the transformations. In those plant strains where *Agrobacterium*-mediated transformation is efficient, it is the method of choice because of the facile and defined nature of the gene transfer.

A transgenic plant formed using *Agrobacterium* transformation methods typically contains a single gene on one chromosome. Such transgenic plants can be referred to as being heterozygous for the added gene. More preferred is a transgenic plant that is homozygous for the added structural gene; *i.e.*, a transgenic plant that contains two added genes, one gene at the same locus on each chromosome of a chromosome pair. A homozygous transgenic plant can be obtained by sexually mating (selfing) an independent segregant transgenic plant that contains a single added gene, germinating some of the seed produced and analyzing the resulting plants produced for the gene of interest.

It is also to be understood that two different transgenic plants can also be mated to produce offspring that contain two independently segregating, exogenous genes. Selfing of appropriate progeny can produce plants that are homozygous for both added, exogenous genes that encode a polypeptide of interest. Backcrossing to a parental plant

and out-crossing with a non-transgenic plant are also contemplated, as is vegetative propagation.

Transformation of plant protoplasts can be achieved using methods based on calcium phosphate precipitation, polyethylene glycol treatment, electroporation and combinations of these treatments (*See, for example, Potrykus et al., Mol. Gen. Genet.* 205:193-200 (1986); Lorz *et al., Mol. Gen. Genet.* 199:178 (1985); Fromm *et al., Nature* 319:791 (1986); Uchimiya *et al., Mol. Gen. Genet.* 204:204 (1986); Marcotte *et al., Nature* 335:454-457 (1988)).

Application of these systems to different plant strains depends upon the ability to regenerate that particular plant strain from protoplasts. Illustrative methods for the regeneration of cereals from protoplasts are described (Fujimura *et al., Plant Tissue Culture Letters* 2:74 (1985); Toriyama *et al., Theor Appl. Genet.* 205:34 (1986); Yamada *et al., Plant Cell Rep.* 4:85 (1986); Abdullah *et al., Biotechnology* 4:1087 (1986)).

To transform plant strains that cannot be successfully regenerated from protoplasts, other ways to introduce DNA into intact cells or tissues can be utilized. For example, regeneration of cereals from immature embryos or explants can be effected as described (Vasil, *Biotechnology* 6:397 (1988)). In addition, "particle gun" or high-velocity microprojectile technology can be utilized (Vasil *et al., Bio/Technology* 10:667 (1992)).

Using the latter technology, DNA is carried through the cell wall and into the cytoplasm on the surface of small metal particles as described (Klein *et al., Nature* 328:70 (1987); Klein *et al., Proc. Natl. Acad. Sci. (U.S.A.)* 85:8502-8505 (1988); McCabe *et al., Bio/Technology* 6:923 (1988)). The metal particles penetrate through several layers of cells and thus allow the transformation of cells within tissue explants.

The regeneration, development and cultivation of plants from single plant protoplast transformants or from various transformed explants are well known in the art (Weissbach *et al., In: Methods for Plant Molecular Biology*, Academic Press, San Diego, CA, (1988)). This regeneration and growth process typically includes the steps of selection of transformed cells, culturing those individualized cells through the usual

stages of embryonic development through the rooted plantlet stage. Transgenic embryos and seeds are similarly regenerated. The resulting transgenic rooted shoots are thereafter planted in an appropriate plant growth medium such as soil.

The development or regeneration of plants containing the foreign, exogenous gene that encodes a protein of interest is well known in the art. Preferably, the regenerated plants are self-pollinated to provide homozygous transgenic plants. Otherwise, pollen obtained from the regenerated plants is crossed to seed-grown plants of agronomically important lines. Conversely, pollen from plants of these important lines is used to pollinate regenerated plants. A transgenic plant of the invention containing a desired polypeptide is cultivated using methods well known to one skilled in the art.

There are a variety of methods for the regeneration of plants from plant tissue. The particular method of regeneration will depend on the starting plant tissue and the particular plant species to be regenerated.

Methods for transforming dicots, primarily by use of *Agrobacterium tumefaciens* and obtaining transgenic plants have been published for cotton (U.S. Patent No. 5,004,863; U.S. Patent No. 5,159,135; U.S. Patent No. 5,518,908); soybean (U.S. Patent No. 5,569,834; U.S. Patent No. 5,416,011; McCabe *et al.*, *Biotechnology* 6:923 (1988); Christou *et al.*, *Plant Physiol.* 87:671-674 (1988)); *Brassica* (U.S. Patent No. 5,463,174); peanut (Cheng *et al.*, *Plant Cell Rep.* 15:653-657 (1996), McKently *et al.*, *Plant Cell Rep.* 14:699-703 (1995)); papaya; and pea (Grant *et al.*, *Plant Cell Rep.* 15:254-258 (1995)).

Transformation of monocotyledons using electroporation, particle bombardment and *Agrobacterium* have also been reported. Transformation and plant regeneration have been achieved in asparagus (Bytebier *et al.*, *Proc. Natl. Acad. Sci. (USA)* 84:5354 (1987)); barley (Wan *et al.*, *Plant Physiol* 104:37 (1994)); maize (Rhodes *et al.*, *Science* 240:204 (1988); Gordon-Kamm *et al.*, *Plant Cell* 2:603-618 (1990); Fromm *et al.*, *Bio/Technology* 8:833 (1990); Koziel *et al.*, *Bio/Technology* 11:194 (1993); Armstrong *et al.*, *Crop Science* 35:550-557 (1995)); oat (Somers *et al.*, *Bio/Technology* 10:1589 (1992)); orchard grass (Horn *et al.*, *Plant Cell Rep.* 7:469 (1988)); rice (Toriyama *et al.*, *Theor Appl. Genet.* 205:34 (1986); Part *et al.*, *Plant Mol. Biol.* 32:1135-1148 (1996);

Abedinia *et al.*, *Aust. J. Plant Physiol.* 24:133-141 (1997); Zhang *et al.*, *Theor. Appl. Genet.* 76:835 (1988); Zhang *et al.*, *Plant Cell Rep.* 7:379 (1988); Battraw *et al.*, *Plant Sci.* 86:191-202 (1992); Christou *et al.*, *Bio/Technology* 9:957 (1991)); rye (De la Pena *et al.*, *Nature* 325:274 (1987)); sugarcane (Bower *et al.*, *Plant J.* 2:409 (1992)); tall fescue (Wang *et al.*, *Bio/Technology* 10:691 (1992)) and wheat (Vasil *et al.*, *Bio/Technology* 10:667 (1992); U.S. Patent No. 5,631,152).

Assays for gene expression based on the transient expression of cloned nucleic acid constructs have been developed by introducing the nucleic acid molecules into plant cells by polyethylene glycol treatment, electroporation, or particle bombardment (Marcotte *et al.*, *Nature* 335:454-457 (1988); Marcotte *et al.*, *Plant Cell* 1:523-532 (1989); McCarty *et al.*, *Cell* 66:895-905 (1991); Hattori *et al.*, *Genes Dev.* 6:609-618 (1992); Goff *et al.*, *EMBO J.* 9:2517-2522 (1990)). Transient expression systems may be used to functionally dissect gene constructs (*see generally*, Mailga *et al.*, *Methods in Plant Molecular Biology*, Cold Spring Harbor Press (1995)).

Any of the nucleic acid molecules of the invention may be introduced into a plant cell in a permanent or transient manner in combination with other genetic elements such as vectors, promoters, enhancers, *etc.* Further, any of the nucleic acid molecules of the invention may be introduced into a plant cell in a manner that allows for overexpression of the protein or fragment thereof encoded by the nucleic acid molecule.

Cosuppression is the reduction in expression levels, usually at the level of RNA, of a particular endogenous gene or gene family by the expression of a homologous sense construct that is capable of transcribing mRNA of the same strandedness as the transcript of the endogenous gene (Napoli *et al.*, *Plant Cell* 2:279-289 (1990); van der Krol *et al.*, *Plant Cell* 2:291-299 (1990)). Cosuppression may result from stable transformation with a single copy nucleic acid molecule that is homologous to a nucleic acid sequence found within the cell (Prolls *et al.*, *Plant J.* 2:465-475 (1992)) or with multiple copies of a nucleic acid molecule that is homologous to a nucleic acid sequence found within the cell (Mittlesten *et al.*, *Mol. Gen. Genet.* 244:325-330 (1994)). Genes, even though different, linked to homologous promoters may result in the cosuppression of the linked genes

(Vaucheret, *C.R. Acad. Sci. III* 316:1471-1483 (1993); Flavell, *Proc. Natl. Acad. Sci. (U.S.A.)* 91:3490-3496 (1994)); van Blokland *et al.*, *Plant J.* 6:861-877 (1994); Jorgensen, *Trends Biotechnol.* 8:340-344 (1990); Meins *et al.*, In: *Gene Inactivation and Homologous Recombination in Plants*, Paszkowski (ed.), pp. 335-348, Kluwer Academic, Netherlands (1994)).

It is understood that one or more of the nucleic acids of the invention may be introduced into a plant cell and transcribed using an appropriate promoter with such transcription resulting in the cosuppression of an endogenous protein.

Antisense approaches are a way of preventing or reducing gene function by targeting the genetic material (U.S. Patents 4,801,540 and 5,107,065 Mol *et al.*, *FEBS Lett.* 268:427-430 (1990)). The objective of the antisense approach is to use a sequence complementary to the target gene to block its expression and create a mutant cell line or organism in which the level of a single chosen protein is selectively reduced or abolished. Antisense techniques have several advantages over other 'reverse genetic' approaches. The site of inactivation and its developmental effect can be manipulated by the choice of promoter for antisense genes or by the timing of external application or microinjection. Antisense can manipulate its specificity by selecting either unique regions of the target gene or regions where it shares homology to other related genes (Hiatt *et al.*, In: *Genetic Engineering*, Setlow (ed.), Vol. 11, New York: Plenum 49-63 (1989)).

The principle of regulation by antisense RNA is that RNA that is complementary to the target mRNA is introduced into cells, resulting in specific RNA:RNA duplexes being formed by base pairing between the antisense substrate and the target mRNA (Green *et al.*, *Annu. Rev. Biochem.* 55:569-597 (1986)). Under one embodiment, the process involves the introduction and expression of an antisense gene sequence. Such a sequence is one in which part or all of the normal gene sequences are placed under a promoter in inverted orientation so that the 'wrong' or complementary strand is transcribed into a noncoding antisense RNA that hybridizes with the target mRNA and interferes with its expression (Takayama *et al.*, *Crit. Rev. Biochem. Mol. Biol.* 25:155-184 (1990)). An antisense vector is constructed by standard procedures and introduced into

cells by transformation, transfection, electroporation, microinjection, infection, *etc.* The type of transformation and choice of vector will determine whether expression is transient or stable. The promoter used for the antisense gene may influence the level, timing, tissue, specificity, or inducibility of the antisense inhibition.

It is understood that the activity of a protein in a plant cell may be reduced or depressed by growing a transformed plant cell containing a nucleic acid molecule whose non-transcribed strand encodes a protein or fragment thereof.

Post transcriptional gene silencing (PTGS) can result in virus immunity or gene silencing in plants. PTGS is induced by dsRNA and is mediated by an RNA-dependent RNA polymerase, present in the cytoplasm, that requires a dsRNA template. The dsRNA is formed by hybridization of complementary transgene mRNAs or complementary regions of the same transcript. Duplex formation can be accomplished by using transcripts from one sense gene and one antisense gene co-located in the plant genome, a single transcript that has self-complementarity, or sense and antisense transcripts from genes brought together by crossing. The dsRNA-dependent RNA polymerase makes a complementary strand from the transgene mRNA and RNase molecules attach to this complementary strand (cRNA). These cRNA-RNase molecules hybridize to the endogene mRNA and cleave the single-stranded RNA adjacent to the hybrid. The cleaved single-stranded RNAs are further degraded by other host RNases because one will lack a capped 5' end and the other will lack a poly(A) tail (Waterhouse *et al.*, *PNAS* 95: 13959-13964 (1998)).

It is understood that one or more of the nucleic acids of the invention may be introduced into a plant cell and transcribed using an appropriate promoter with such transcription resulting in the postranscriptional gene silencing of an endogenous transcript.

Antibodies have been expressed in plants (Hiatt *et al.*, *Nature* 342:76-78 (1989); Conrad *et al.*, *Plant Mol. Biol.* 26:1023-1030 (1994)). Cytoplasmic expression of a scFv (single-chain Fv antibodies) has been reported to delay infection by artichoke mottled crinkle virus. Transgenic plants that express antibodies directed against endogenous

proteins may exhibit a physiological effect (Philips *et al.*, *EMBO J.* 16:4489-4496 (1997); Marion-Poll, *Trends in Plant Science* 2:447-448 (1997)). For example, expressed anti-abscissic antibodies have been reported to result in a general perturbation of seed development (Philips *et al.*, *EMBO J.* 16: 4489-4496 (1997)).

Antibodies that are catalytic may also be expressed in plants (abzymes). The principle behind abzymes is that since antibodies may be raised against many molecules, this recognition ability can be directed toward generating antibodies that bind transition states to force a chemical reaction forward (Persidas, *Nature Biotechnology* 15:1313-1315 (1997); Baca *et al.*, *Ann. Rev. Biophys. Biomol. Struct.* 26:461-493 (1997)). The catalytic abilities of abzymes may be enhanced by site directed mutagenesis. Examples of abzymes are, for example, set forth in U.S. Patent No. 5,658,753; U.S. Patent No. 5,632,990; U.S. Patent No. 5,631,137; U.S. Patent 5,602,015; U.S. Patent No. 5,559,538; U.S. Patent No. 5,576,174; U.S. Patent No. 5,500,358; U.S. Patent 5,318,897; U.S. Patent No. 5,298,409; U.S. Patent No. 5,258,289 and U.S. Patent No. 5,194,585.

It is understood that any of the antibodies of the invention may be expressed in plants and that such expression can result in a physiological effect. It is also understood that any of the expressed antibodies may be catalytic.

#### **(d) Antibodies**

One aspect of the present invention concerns antibodies, single-chain antigen binding molecules, or other proteins that specifically bind to one or more of the protein or peptide molecules of the present invention and their homologues, fusions or fragments. Such antibodies may be used to quantitatively or qualitatively detect the protein or peptide molecules of the present invention. As used herein, an antibody or peptide is said to "specifically bind" to a protein or peptide molecule of the present invention if such binding is not competitively inhibited by the presence of non-related molecules.

Nucleic acid molecules that encode all or part of the protein of the present invention can be expressed, via recombinant means, to yield protein or peptides that can in turn be used to elicit antibodies that are capable of binding the expressed protein or peptide. Such antibodies may be used in immunoassays for that protein. Such protein-



encoding molecules, or their fragments may be a "fusion" molecule (*i.e.*, a part of a larger nucleic acid molecule) such that, upon expression, a fusion protein is produced. It is understood that any of the nucleic acid molecules of the present invention may be expressed, via recombinant means, to yield proteins or peptides encoded by these nucleic acid molecules.

The antibodies that specifically bind proteins and protein fragments of the present invention may be polyclonal or monoclonal and may comprise intact immunoglobulins, or antigen binding portions of immunoglobulins fragments (such as  $F(ab')$ ,  $F(ab')_2$ ), or single-chain immunoglobulins producible, for example, via recombinant means. It is understood that practitioners are familiar with the standard resource materials which describe specific conditions and procedures for the construction, manipulation and isolation of antibodies (*see*, for example, Harlow *et al.*, In: *Antibodies: A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring Harbor, New York (1988)).

Murine monoclonal antibodies are particularly preferred. BALB/c mice are preferred for this purpose, however, equivalent strains may also be used. The animals are preferably immunized with approximately 25  $\mu$ g of purified protein (or fragment thereof) that has been emulsified in a suitable adjuvant (such as TiterMax adjuvant (Vaxcel, Norcross, GA)). Immunization is preferably conducted at two intramuscular sites, one intraperitoneal site and one subcutaneous site at the base of the tail. An additional *i.v.* injection of approximately 25  $\mu$ g of antigen is preferably given in normal saline three weeks later. After approximately 11 days following the second injection, the mice may be bled and the blood screened for the presence of anti-protein or peptide antibodies. Preferably, a direct binding Enzyme-Linked Immunoassay (ELISA) is employed for this purpose.

More preferably, the mouse having the highest antibody titer is given a third *i.v.* injection of approximately 25  $\mu$ g of the same protein or fragment. The splenic leukocytes from this animal may be recovered 3 days later and then permitted to fuse, most preferably, using polyethylene glycol, with cells of a suitable myeloma cell line (such as, for example, the P3X63Ag8.653 myeloma cell line). Hybridoma cells are selected by

culturing the cells under "HAT" (hypoxanthine-aminopterin-thymine) selection for about one week. The resulting clones may then be screened for their capacity to produce monoclonal antibodies ("mAbs"), preferably by direct ELISA.

In one embodiment, anti-protein or peptide monoclonal antibodies are isolated using a fusion of a protein or peptide of the present invention, or conjugate of a protein or peptide of the present invention, as immunogens. Thus, for example, a group of mice can be immunized using a fusion protein emulsified in Freund's complete adjuvant (*e.g.*, approximately 50 µg of antigen per immunization). At three week intervals, an identical amount of antigen is emulsified in Freund's incomplete adjuvant and used to immunize the animals. Ten days following the third immunization, serum samples are taken and evaluated for the presence of antibody. If antibody titers are too low, a fourth booster can be employed. Polysera capable of binding the protein or peptide can also be obtained using this method.

In a preferred procedure for obtaining monoclonal antibodies, the spleens of the above-described immunized mice are removed, disrupted and immune splenocytes are isolated over a ficoll gradient. The isolated splenocytes are fused, using polyethylene glycol with BALB/c-derived HGPRT (hypoxanthine guanine phosphoribosyl transferase) deficient P3x63xAg8.653 plasmacytoma cells. The fused cells are plated into 96 well microtiter plates and screened for hybridoma fusion cells by their capacity to grow in culture medium supplemented with hypoxanthine, aminopterin and thymidine for approximately 2-3 weeks.

Hybridoma cells that arise from such incubation are preferably screened for their capacity to produce an immunoglobulin that binds to a protein of interest. An indirect ELISA may be used for this purpose. In brief, the supernatants of hybridomas are incubated in microtiter wells that contain immobilized protein. After washing, the titer of bound immunoglobulin can be determined using, for example, a goat anti-mouse antibody conjugated to horseradish peroxidase. After additional washing, the amount of immobilized enzyme is determined (for example through the use of a chromogenic substrate). Such screening is performed as quickly as possible after the identification of

the hybridoma in order to ensure that a desired clone is not overgrown by non-secreting neighbor cells. Desirably, the fusion plates are screened several times since the rates of hybridoma growth vary. In a preferred sub-embodiment, a different antigenic form may be used to screen the hybridoma. Thus, for example, the splenocytes may be immunized with one immunogen, but the resulting hybridomas can be screened using a different immunogen. It is understood that any of the protein or peptide molecules of the present invention may be used to raise antibodies.

Such antibody molecules or their fragments may be used for diagnostic purposes. Where the antibodies are intended for diagnostic purposes, it may be desirable to derivatize them, for example with a ligand group (such as biotin) or a detectable marker group (such as a fluorescent group, a radioisotope or an enzyme).

The ability to produce antibodies that bind the protein or peptide molecules of the present invention permits the identification of mimetic compounds of those molecules. A "mimetic compound" is a compound that is not that compound, or a fragment of that compound, but which nonetheless exhibits an ability to specifically bind to antibodies directed against that compound.

Having now generally described the invention, the same will be more readily understood through reference to the following examples which are provided by way of illustration, and are not intended to be limiting of the present invention, unless specified.

#### EXAMPLE 1

In this example, DNA is extracted from soybean plants, amplified, and mapped.

A single trifoliate leaf is collected from the newest growth of four week old soybean plants. Leaf tissue from the leaf is placed on ice and stored at  $-80^{\circ}\text{C}$ . The frozen tissue is lyophilized, and approximately 0.01 grams of the tissue is used for DNA extraction. The 0.01 grams of leaf tissue is ground to powder in 1.4 ml tubes. 600 microliters ( $\mu\text{l}$ ) of DNA extraction buffer consisting of 0.5M NaCl, 0.1M Tris-(hydroxymethyl) aminomethane pH 8.0, 0.05 M ethylenediaminetetra-acetic acid (EDTA),  $10.0 \text{ g L}^{-1}$  sodium dodecyl sulfate (SDS), and  $2 \text{ g L}^{-1}$  phenantroline (dissolved in 0.01 L ethanol) is heated to  $65^{\circ}\text{C}$  (with  $0.77 \text{ g L}^{-1}$  dithiothreitol added immediately before

use) is added to each tube, and each tube is mixed thoroughly. The samples are placed in a 65°C water bath for 15 minutes and shaken by hand after 10 minutes. The samples are taken out of the water bath and cooled to room temperature, and then 200 µl of 5 M KOAc is added to each tube. The samples are inverted and placed at 4°C for 20 minutes. Samples are then centrifuged for 12 minutes at 6200 X g and the supernatant (about 600 µl) is transferred to new tubes. DNA is precipitated with 330 µl of cold isopropanol and placed at -20°C for 1 hr. The DNA is pelleted by centrifuging at 6200 X g for 10 minutes and washed with 70% EtOH. The DNA is pelleted by centrifugation at 6200 X g for 10 minutes and dried using a Speed-Vac. The DNA is dissolved in 100 µl of TE<sub>0.1</sub> (0.01 M Tris-HCl pH 8.0, 0.0001 M EDTA). The extraction will generally yield 500 ng DNA µl<sup>-1</sup>.

A polymerase chain reaction (PCR) is conducted with 5 to 10 ng genomic DNA in 10 µl volumes of 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 0.001% gelatin, 1.5 mM MgCl<sub>2</sub>, 0.1 mM of each dNTP, 150 nM of each primer, 0.01 mM Cresol Red, 2% sucrose and 0.32 units of AmpliTaq DNA Polymerase (Perkin Elmer Instruments Inc., USA). For thermocycling, the Gene Amp PCR System 9700 (Perkin Elmer Instruments Inc., USA) is used with one step of 94°C for 3 minutes, then 32 cycles of 94°C, 47°C, and 72°C steps of 25 sec each and one final step of 72°C for 3 minutes. The PCR products are run on a 6% polyacrylamide gel (30 cm X 8 cm X 1 mm) in 1X TAE (40 mM Tris-HCl, pH 8.3, 1 mM EDTA) at 180 v for 45 minutes. The gels are stained using SYBR Gold (Molecular Probes, Eugene, OR) according to the manufacturer's instructions.

SSR primer screening for polymorphism is performed using PIC, HS-1, Will and PI507354 genotypes. SSRs that are polymorphic and easy to score (*i.e.*, clear banding pattern and good separation between alleles) are mapped using the HS-1 x PIC (F2) and/or Will x PI507354 (RIL) mapping populations. At least one SSR per BAC sequence is mapped. DNA markers that exhibited codominant banding patterns are scored as homozygous for one or the other parent or as heterozygous, exhibiting both parental alleles. Marker scores are checked for segregation distortion using the chi-squared test for goodness of fit to expected ratios. Linkage relationships are determined using Mapmaker Version 3.0b with a LOD of 3.0 (Whitehead Institute, Cambridge, MA).

## EXAMPLE 2

DNA fragments containing candidates for genes *rhg1* and *Rhg4* from susceptible and resistant soybean lines are subcloned into a TA cloning plasmid (TOPO TA Cloning Kit, Version E, Invitrogen Corporation, 1600 Faraday Avenue, Carlsbad, CA).

Genomic DNA from 24 susceptible and 9 resistant lines is isolated using standard techniques. Approximately 500 nanograms (ng) of DNA is used for PCR amplification. Resistant BAC DNA is isolated by using AUTOGEN (AutoGen Corp., 35 Loring Drive Framingham, MA). PCR amplification is then performed using 0.1- 0.2 ng of resistant BAC DNA. The primers that are used to amplify candidate *rhg1* genes PCR are as follows:

Fragment I (2,892 bp) primer (SEQ ID NO: 25), GCA ATA CTT GAA GGA ATA TGT CCA C; primer (SEQ ID NO: 24), beginning at start codon, ATG GAT GGT AAA AAT TCA AAA CTA AAC; modified reverse primer 1 (SEQ ID NO: 1123), beginning 5 bp before start codon; GTT GTA TGG ATG GTA AAA ATT CAA AAC. Fragment II (1,746 bp) reverse primer 2 (SEQ ID NO: 27), ending at 13 bp after stop codon, GAC TGG CTG TGA CTG ATC TCT CT; primer 2 (SEQ ID NO: 26), CTC ACT TAC ACT GCT GAA TGC AGA.

The primers for *Rgh4* PCR are as follows:

Forward primer (SEQ ID NO: 48), ATG TCT CTC CCC AAA ACC CTA CTT TCT CTC; reverse primer (SEQ ID NO: 49), ending at 2 bp after stop codon, GGT TAA CGG CAA TCC ATT GAA TCA AAG GAG.

PCR amplification is performed in an MJ Research PTC DNA Engine *TM* System, Model PTC-225 (MJ Research Inc, 590 Lincoln Street Waltham, MA). PCR is performed using the following components: 1µl DNA, 5µl 10x buffer, 1µl primer 1, 1µl primer 2, 1µl 10mM dNTP, 1.5 µl 50mM MgCl<sub>2</sub>, 0.2µl Taq. (Platinum), 39.3µl H<sub>2</sub>O. The PCR program used is as follows: 95°C for 10 minutes (step 1), 95°C for 30 seconds (step 2), 70°C for 30 seconds/-1°C per cycle/72°C for 3 minutes (step 3), repeat steps two through three 9 times (step 4), 95°C for 30 seconds (step 5), 60°C for 30 seconds (step 6),

72°C for 3 minutes (step 7), repeat steps five through seven 34 times (step 8), 4°C forever (step 9), end.

PCR products are separated on 1% agarose gel by electrophoresis. A single DNA band is excised from gel. Gel extraction is done using CLONTECH NucleoSpin Extraction Kit (Clontech Laboratories Inc., 1020 East Meadow Circle, Palo Alto, CA). 2 µl of purified DNA is loaded on 1% agarose gel to check concentration. 40-100ng of DNA is used for subcloning.

A TOPO cloning reaction is done according to the following: 4µl of fresh PCR product, 1µl Clontech Salt Solution, and 1 µl TOPO vector. The solution is mixed gently, incubated for 10 minutes at room temperature, and then placed on ice.

A one shot chemical transformation is performed as follows. 2µl of the TOPO Cloning reaction is added to a vial of TOP 10 One Shot Chemically Competent *E. coli* and mixed gently. The mixture is then Incubated on ice for 30 minutes. The cells are then heat-shocked for 30 seconds at 42°C, and immediately transferred to ice. 250 µl of SOC medium is then added, and the mixture is incubated at 37°C for 1 hour. 80 µl is then spread onto a selective plate, and 170 µl is spread onto another plate. The plates are incubated at 37°C for 18-20 hours. The selective plates are LB agar plates with 100 µg/ml ampicillin, 40 µg/ml IPTG, and 40 µg/ml X-GAL.

After incubation, 8-10 white or light blue colonies are selected. The positive colonies are inoculated into LB medium containing 50 µg/ml ampicillin and incubated at 37°C overnight. Sterilized glycerol is added to make 15% glycerol stock, which can be stored at -80°C.

Sanger sequencing reactions are performed on subclones using BigDye Terminators (Applied Biosystems, 850 Lincoln Centre Drive, Foster City, CA) and then analyzed on ABI 377/ABI 3700 automated sequencing machines (Applied Biosystems, 850 Lincoln Centre Drive, Foster City, CA). The sequences are evaluated for quality and error probability using the program, PHRED (Ewing and Green, *Genome Res.*, 8:186-194 (1998), Ewing *et al.*, *Genome Res.*, 8:175-185, (1998)), assembled using the *phrap* assembler and viewed using *consed* (Gordon *et al.*, *Genome Res.*, 8:195-202). An *rhg1*

candidate gene is found in BAC 240O17, and is about 4.5 kb in size. An *Rhg4* candidate was found in BAC 318O13, and is about 3.5 kb in size.

### EXAMPLE 3

The physical mapping of a QTL (quantitative trait locus) is described in this example. Mapping is initiated with linkage analysis of SSR (simple sequence repeats) markers. Markers that are shown to be linked to the QTL of interest are used to PCR screen the soy BAC library and identify candidate BACs. Confirmed BACs are subcloned and sequenced, BAC-end sequenced, and fingerprinted. New markers are designed from good BAC-end sequences and used to screen the library, by either PCR or hybridization to high density grid filters, in order to extend the contigs. A BAC-end sequence and fingerprint database of soy BACs is used in conjunction with the above methods to help build and extend contigs. Sequenced BACs are aligned, and overlapping BACs are placed into contigs. These contigs, which contain unique sequences, are put into an ACEDB database, and predicted genes are annotated by hand using various programs. Candidates genes (for the gene of interest) are subcloned from genomic DNA of different lines by PCR using primers from outside the predicted coding regions. These subclones are sequenced and screened for SNPs (single nucleotide polymorphisms) and INDELs (insertions/deletions), and different haplotypes of the lines with and without the desired phenotype are examined for correlations between the haplotype and phenotype.

A single trifoliate leaf is collected from the newest growth of four week old soybean plants. The leaf tissue is placed on ice and stored at -80°C. The frozen tissue is lyophilized and approximately 0.01 grams of tissue is used for DNA extraction. The leaf tissue is ground to powder in 1.4 ml tubes and 600 µl of DNA extraction buffer [0.5M NaCl, 0.1M Tris-(hydroxymethyl) aminomethane pH 8.0, 0.05 M ethylenediaminetetraacetic acid (EDTA), 10.0 g L<sup>-1</sup> sodium dodecyl sulfate (SDS), 2 g L<sup>-1</sup> phenantroline (dissolved in 0.01 L ethanol)] heated to 65°C (with 0.77 g L<sup>-1</sup> dithiothreitol added immediately before use) is added to each tube and mixed thoroughly. The samples are placed in a 65°C water bath for 15 minutes and shaken by hand after 10 min. The samples are taken out of the water bath, cooled to room temperature, and 200 µl of 5 M

KOAc is added to each tube. The samples are inverted and placed at 4°C for 20 min. Samples are then centrifuged for 12 minutes at 6200 X g and the supernatant (about 600 µl) is transferred to new tubes. DNA is precipitated with 330 µl of cold isopropanol and placed at -20°C for 1 hr. The DNA is pelleted by centrifuging at 6200 X g for 10 minutes and is washed with 70% EtOH. The DNA is pelleted by centrifugation at 6200 X g for 10 minutes and dried using a Speed-Vac. The DNA is dissolved in 100 µl of TE<sub>0.1</sub> (0.01 M Tris-HCl pH 8.0, 0.0001 M EDTA). The extraction yields 500 ng DNA µl<sup>-1</sup>.

The polymerase chain reaction (PCR) is conducted with 5 to 10 ng genomic DNA in 10 µl volumes of 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 0.001% gelatin, 1.5 mM MgCl<sub>2</sub>, 0.1 mM of each dNTP, 150 nM of each primer, 0.01 mM Cresol Red, 2% sucrose and 0.32 units of AmpliTaq DNA Polymerase (Perkin Elmer Instruments Inc., USA, 761 Main Avenue, Norwalk, CT). For thermocycling, the Gene Amp PCR System 9700 (Perkin Elmer Instruments Inc., USA, 761 Main Avenue, Norwalk, CT) is used with one step of 94°C for 3 min, then 32 cycles of 94°C, 47°C and 72°C steps of 25 sec each and one final step of 72°C for 3 min. The PCR products are run on a 6% polyacrylamide gel (30 cm X 8 cm X 1 mm) in 1X TAE (40 mM Tris-HCl, pH 8.3, 1 mM EDTA) at 180v for 45 min. The gels are stained using SYBR Gold (Molecular Probes, Eugene, OR) per manufacturers instructions.

SSR primer screening for polymorphisms is performed using PIC, HS-1, Will and PI507354 genotypes. SSRs that are polymorphic and easy to score (*i.e.*, Clear banding pattern and good separation between alleles) are mapped using the HS-1 x PIC (F2) and/or Will x PI507354 (RIL) mapping populations. At least one SSR per BAC sequence is mapped. DNA markers that exhibited codominant banding patterns are scored as homozygous for one or the other parent or as heterozygous, exhibiting both parental alleles. Marker scores are checked for segregation distortion using the chi-squared test for goodness of fit to expected ratios. Linkage relationships were determined using Mapmaker Version 3.0b with a LOD of 3.0 (Whitehead Institute for Biomedical Research, Cambridge MA).



Thirty-two BAC DNA superpools (10 genomic equivalents) extracted from either 4608 clones (48 96-well microtiter plates) are used as templates for the first round of PCR screening. Following identification of the positive superpools, the second screening is performed against 4-D BAC DNA pools. Each clone of the superpool is addressed 4-dimensionally (7 X 7 X 12 X 8) and pooled in each dimension. Each set of 48 plates is divided into 6 sets of 7 plates and one set of 6 plates, and partitioned in two ways. The first partition is in numerical order, plates 1-7, 8-14, ... 43-48 representing 7 group or stack pools. The second partition is according to plate position within each of the respective stacks, plates [1, 8, 15, 22, 29, 36], [2,9,16, 23, 30, 37, 43] *etc.*, representing 7 plate pools. Each well of the 96-well plates contains 12 columns and 8 rows. Clones from row 1 are pooled from all 48 plates to generate the row 1 pool. Clones of rows 2, 3, 4,...8, and columns 1, 2, 3,...12 are pooled to generate 8 row pools and 12 column pools respectively.

For each superpool, BAC DNA is extracted from a total of 34 subpools (7 + 7 + 8 + 12). Positive clones are identified by TaqMan/PCR screening of the 34 subpools if one positive clone is present. If more than one positive clone is present in a superpool, a third round of screening with N4 PCR reactions is performed.

Addresses of candidate BACs are identified, and the candidates are streaked out for single colony isolation and grown overnight at 37°C. A single, isolated colony is picked and streaked out and grown overnight at 37°C. PCR is repeated for the marker of interest (using the program designed for the relevant marker) using a smear of cells from the plate streaked from a single colony. The PCR product is run on a 2% agarose gel and purified using the Clontech NucleoSpin Gel Extraction Kit (according to the manufacturer's instructions, Clontech Laboratories Inc., 1020 East Meadow Circle, Palo Alto, CA) and 10-50 ng of the purified DNA are added to 10 pmol of each primer (forward and reverse), in a total volume of 6 µl of ddH<sub>2</sub>O and 2 µl of BigDye Terminators (Applied Biosystems, 850 Lincoln Centre Drive, Foster City, CA). The cycling conditions are: 96°C for 1 minute (step 1), 96°C for 10 seconds (step 2), 50°C for

5 seconds (step 3), 60°C for 4 minutes (step 4), steps 2-4 are repeated for 24 cycles (step 5), and hold at 4°C.

The generated sequence is compared to the consensus sequence using DNA comparison software. Confirmed clones are subcloned, sequenced, BAC-end sequenced, and Fingerprinted.

BAC-end sequencing is done using 3.2 pmol of SP6 and T7 primers (separately), approximately 600 ng-1 ug of BAC DNA (Autogen prepped, AutoGen Corp., 35 Loring Drive Framingham, MA) reaction, resuspended in 6µl of ddH<sub>2</sub>O, and 4µl of BigDye Terminators (Applied Biosystems 850 Lincoln Centre Drive, Foster City, CA) to give a total reaction volume of 10ul. The cycling conditions are: 96°C for 2 minutes (step 1), 96°C for 15 seconds (step 2), 50°C for 15 seconds (step 3), 60°C for 4 minutes (step 4), steps 2-4 are repeated for 50-60 cycles (step 5), 72°C for 2 minutes (step 6), hold at 4°C or 10°C (step 7).

The reactions are ethanol precipitated and loaded on capillary sequencers. The newly generated BAC-end sequence is trimmed from the vector sequence, and entered into a database containing approximately 400,000 BAC-end sequences. Each BAC is blasted against the database to search for BAC-end matches extension of the contigs. New markers are designed from good BAC-end sequences, and these are then used to rescreen the library in order to build up contigs across the region of interest. Screening can be done in either of two ways: as above (PCR strategy), or by hybridization of high-density grid filters from Research Genetics (Research Genetics, 2130 Memorial Parkway, Huntsville, AL).

The probes used for hybridization are derived from clones or genomic DNA by PCR amplification using the vector or gene-specific primers, with the appropriate cycling conditions. PCR products are run on a 1% agarose gel containing ethidium bromide (0.2 ug/ml) in 1X TAE buffer at 100 volt for 1-2 hrs. Isolated DNA fragments are excised and gel-purified using the Clonotech NucleoSpin gel extraction kit (Clonotech Laboratories Inc., 1020 East Meadow Circle, Palo Alto, CA), before labeling. In order to check the size of the fragments and concentration, 2 µl of eluted DNA plus loading buffer are

loaded on a 1% agarose gel along with DNA markers of known concentration and size. All the probes used to screen the library are tested individually for repetitiveness, with a smaller filter spotted with random clones from the library along with some positive control clones according to the protocol described below.

The A3244 soy library generated by a an *EcoRI* digest is spotted on 3 high density grid filters from Research Genetics (Research Genetics, 2130 Memorial Parkway, Huntsville, AL). Each filter has six fields, twelve 384 well plates are spotted in each field in duplicate, with a total of 27,648 clones spotted on each filter. The plates are spotted in a 5X5 grid (12 clones per 5X5 grid) pattern within each field. Each clone is spotted in duplicate with a specific orientation within the 5X5 grid, which, together with the field position, gives information about its address. In a first round hybridization procedure, multiple probes are labeled separately and then pooled together to hybridize to BAC filters. Positive BACs identified in this procedure are deconvoluted by rehybridization with the individual probes.

A hybridization oven is set at 65°C, and Church Buffer (0.5 M Sodium Phosphate, pH 7.0, 7% SDS, 1% bovine serum albumin, 1 mM EDTA, 100 µg/ml salmon sperm DNA) is prewarmed to 65°C. Membranes are washed in 500 ml of 0.1X SSC, 0.1% SDS in a large container at room temperature for 5 minutes with gentle shaking (50 rpm) on a rotary shaker. The membranes are rinsed with 500 ml of 0.1X SSC (no SDS) for 1 minute. The wash solution is poured off, and 500 ml of 6XSSC (no SDS) is added to equilibrate the membranes. Three filters are placed in a tube. The filters are separated from each other and the sides of the tube by a layer of mesh. Each tube is filled with 6XSSC and shaken gently with the tube vertical to help eliminate bubbles between the filters and tube wall. The 6X SSC solution is poured off, and 25 ml of pre-warmed Church buffer is added. The bottles are rotated in a hybridization oven at 60 rpm and 65°C for 30 minutes or longer.

Probes are labeled using 1 µl of 40-50 uCi/µl [ $\alpha^{32}$ P dCTP], 50 ng of purified DNA in 49 µl of ddH<sub>2</sub>O, and Read-To-Go Labeling Beads from Amersham Pharmacia according to the manufacturers instructions (Amersham Pharmacia, Uppsala, Sweden).

The probes are purified using the Bio-Spin Column P30 from BioRad according to manufacturers instructions (Bio-Rad Laboratories, 3316 Spring Garden Street, Philadelphia, PA). To 1  $\mu$ l of the column-purified probe is added to a minipoly-Q vial (liquid scintillation vial) for each probe. 5 ml of scintillation liquid is added to each vial, and radiation activity for each vial is measured using a liquid scintillation counter.

After the probes are purified and counted for radioactivity, 10-20 probes and one control probe (from 50  $\mu$ l reaction) are pooled with  $10^7$  cpm/probe each, into one 1.5 ml eppendorf tube. The pooled probes are denatured at 99°C in a sand heating block for 10 minutes. The tubes are cooled on ice or ice water about 2 minutes, and then spun down at 14,000 rpm for 30 seconds in microcentrifuge. The tubes are pre-hybridized in 25ml of Church buffer for at least 30 minutes, which is then poured off. 40 ml of fresh hybridization solution (pre-warmed Church buffer) is added. The pooled-probe solution is added to the hybridization tube. The tube is rotated in the hybridization oven at 60 rpm, 65°C overnight.

The probe solution is poured off, 30 ml of pre-warmed (65°C) 1X SSC, 0.1% SDS washing solution is added to the hybridization tube, the hybridization tube is rotated in the hybridization oven (at 65° C) for 15 minutes, and the process is repeat two times. At the last wash, the tube is rotated 180° and at the same speed for 15 minutes at 65°C. The washing solution is poured off, and 2X SSC (no SDS) is added.

Excess liquid is removed from each filter by placing the filter on a piece of 3MM paper. The washed filter is placed on developed film with the DNA-side up (the side BACs were spotted on), covered with Saran wrap, and squeezed to force out liquid and bubbles. The Saran wrap is folded to the other side of the film, fixed it with tape, and then dried Kimwipes. The wrapped filters are placed into a film cassette with the DNA-side up (the side BACs were spotted on), which is placed on BioMax MS film (Biomax Technologies Inc., Vancouver, BC, Canada) in a darkroom, and exposed overnight at room temperature without an intensifying screen. Film is developed with a film developer in the dark room the next day, and each film is labeled with filter number, probe used for hybridization, exposure time, and date.

Starting from Field 3, a 384-well grid is put on the field with the A1 position of the grid on the upper right, and the grid is aligned to the image. The row and column position for each positive clone on the BAC recording spreadsheet is determined and recorded. The pattern of the hybridization signal is matched to known patterns. There are 6 plate reference numbers for each of twelve patterns, which are arranged in the same manner as the 6 fields. Based on the signal pattern and field number, a plate reference number is determined for each positive clone. The grid is moved to the next field and the process is repeated. The original plate number (P) is determined using the following formula:  $P = (N-1) \times 72 + R$ , where N is the filter number on which the identified clone is present and R is the plate reference number previously determined. The complete address of the identified clone is given by the original plate number plus its position on the plate determined previously. BACs' addresses are identified and converted to "imp" files according to a Q-bot file format.

24 working plates are loaded into a Q-bot (Genetix, Queensway, New Milton, Hampshire, United Kingdom) 6-high hotel and media-filled 96-well plates are placed on the deck. The Q-bot is run following the standard manual using the program called "Rearranging98" with the settings given in Appendix III of the accompanying manual: BAC-Picking. Plates containing picked clones are placed in a shaker incubator and grown overnight at 37° C at 200 rpm.

35 µl DNA solution are transferred from 96-well plates into a 384-well plate using a Platemate such that 4 96-well plates of DNA are combined into one 384-well plate. The 384-pin head (puck) is washed in 10% SDS solution for 5 minutes, ultrasonicated in a water bath for 3 minutes, washed with 70% ethanol for 1 min., and air dried for 3 minutes. The 384-well DNA source plates and membranes are arranged on the deck according to the instruction from the manual and the spotted grid design chosen for the membrane. Spotting pattern are designed so that there is one control probe at each of the 4 corners of the membrane. An asymmetric pattern is used to orient filters. The control probe concentration is about 5 ng/ul. Zeus is run according to instructions. If the DNA concentration is lower than 5 ng/ul, the Zeus is run a second time to double the

amount of spotted DNA on the membrane. One of the empty spots is spot dyed, if available, using one 384-well dye plate. If an empty spot is not available, it is printed on one of the DNA spots. This spot marks the position for cutting filters into small membranes (9X12 cm). Membranes are interleaved between 3M papers and left to air-dry. Each corner of each membrane is marked with a permanent marker and numbered. Filters are denatured on the surface of 3M paper soaked with denaturalization solution for 4 minutes, and then neutralized on the surface of 3 M paper soaked with neutralization solution for 5 minutes. The filters are washed with 2XSSC for 5 minutes and then air dried. The filters are then baked at 80°C for 1 hr. and cut into individual small membranes (9X12 cm) according to the marked corner.

To confirm and deconvolute, hybridizations are done as before, but with newly generated filters, and each probe is done separately with a single filter using the smaller tube. 15 ml of Church buffer is used for the hybridization.

Fingerprints are generated by digesting the BAC DNA with Hind III for 3 hours at 37°C and running the reaction on a 0.8% gel at 200V for 19 hours. The gels are stained with SybrGreen, while shaking at room temperature for 45 minutes, and scanned with a FluorImager. The bands are sized using Frag software and the fingerprints are assembled into contigs within FPC. Every time new clones are added the contigs are rebuilt using a tolerance of 10 and a cutoff of  $10^{-9}$ .

Subclones are generated and Sanger sequencing reactions were performed on randomly chosen subclones using BigDye Terminators (Applied Biosystems, 850 Lincoln Centre Drive, Foster City, CA) then analyzed on ABI 377/ABI 3700 automated sequencing machines (Applied Biosystems, 850 Lincoln Centre Drive, Foster City, CA). 7-8 fold sequence coverage is thereby generated across the BAC. The sequences are evaluated for quality and error probability using the program, *phred*, assembled using the *phrap* assembler, and viewed using *consed*, as in example 2. For Bermuda standard BACs, all contigs are ordered and oriented and all gaps are closed using a directed primer walking strategy. A final quality value of *phred40* (1 base error in 10,000 bases) with no gap regions, double coverage or two chemistries across single stranded areas is achieved.

The sequence contigs are put into an ACEDB database along with soy EST and plant EST matches, along with Blastx, Tblastx, and Plant Blastn hits. Genemark.hmm is used to predict possible genes, and GeneFinder is used to predict splicing sites, ORFs, potential coding regions, as well as start and stop codons. The contigs are then annotated by hand and predicted genes accepted, edited, and modified based on the characteristics present in the sequence and matches to protein, nucleotide, and EST databases.

The high-density BAC library membranes used for hybridization are made by Research Genetics (Research Genetics, 2130 Memorial Parkway, Huntsville, AL), using a modified Q-bot (Genetix, Queensway, New Milton, Hampshire, United Kingdom), 384-well plates containing BACs are spotted onto 22 cm X 22 cm Hybond N+ membranes (Amersham Pharmacia, Uppsala, Sweden). Bacteria from 72 plates are spotted twice onto one membrane, giving 55,296 colonies in total, or 27,648 unique clones per membrane. The plates are spotted into six "fields" per membrane, with each field having 12 plates spotted in duplicate. This spotting format results in six fields with 384 grids in each field. Each grid is a 5X5 matrix containing 12 unique clones in duplicate, with the center position left empty. The two positions occupied by each clone in duplicate are designed to give a unique pattern that indicates the plate location of each clone. After spotting, the bacteria on the membrane are incubated for 8 hours on LB-agar plates containing 12.5 ug/ml chloramphenicol. The membranes are then denatured, neutralized, washed in a standard procedure, UV-light crosslinked, and air-dried. The membranes can be stored and shipped at room temperature.

What is claimed is:

1. A method for the production of a soybean plant having an *rhg1* SCN resistant allele comprising:

(A) crossing a first soybean plant having an *rhg1* SCN resistant allele with a second soybean plant having an *rhg1* SCN sensitive allele to produce a segregating population;

(B) screening said segregating population for a member having an *rhg1* SCN resistant allele with a first nucleic acid molecule capable of specifically hybridizing to linkage group G, wherein said first nucleic acid molecule specifically hybridizes to a second nucleic acid molecule that is linked to said *rhg1* SCN resistant allele; and,

(C) selecting said member for further crossing and selection.

2. The method for the production of a soybean plant according to claim 1, wherein said first nucleic acid molecule is capable of specifically hybridizing to said second nucleic acid molecule having the nucleic acid sequence of SEQ ID NO: 2, 3, complements thereof, or fragments thereof having at least 15 nucleotides.

3. The method for the production of a soybean plant according to claim 1, wherein said first nucleic acid molecule is capable of specifically hybridizing to said second nucleic acid molecule having the nucleic acid sequence of SEQ ID NO: 5 or 6, complements thereof, or fragments thereof having at least 15 nucleotides.

4. The method for the production of a soybean plant according to claim 1, wherein said first nucleic acid molecule is a nucleic acid marker capable of detecting *rhg1* haplotype 2 or 4.

5. The method for the production of a soybean plant according to claim 1, wherein said first nucleic acid molecule is capable of specifically hybridizing to a nucleic acid molecule having a sequence that is present on linkage group G within 100kb of said *rhg1* SCN sensitive allele.

6. The method for the production of a soybean plant according to claim 5, wherein said first nucleic acid molecule is capable of specifically hybridizing to a nucleic acid molecule having a sequence that is present on linkage group G and located within 50kb of said *rhg1* SCN sensitive allele.



7. The method for the production of a soybean plant according to claim 6, wherein said first nucleic acid molecule is capable of specifically hybridizing to a nucleic acid molecule having a sequence that is present on linkage group G and located within 25kb of said *rhg1* SCN sensitive allele.
8. The method for the production of a soybean plant according to claim 7, wherein said *rhg1* SCN sensitive allele is also present in soybean line A3244.
9. The method for the production of a soybean plant according to claim 1, wherein said first soybean plant having an *rhg1* SCN resistant allele is selected from the soybean maturity group consisting of 000, 00, 0, I, II, III, IV, and V.
10. The method for the production of a soybean plant according to claim 1, wherein said first soybean plant having an *rhg1* SCN resistant allele is selected from the soybean maturity group consisting of VI, VII, VIII, IX, and X.
11. A method of investigating an *rhg1* haplotype of a soybean plant comprising:
  - (A) isolating nucleic acid molecules from said soybean plant;
  - (B) determining the nucleic acid sequence of an *rhg1* allele or part thereof; and,
  - (C) comparing the nucleic acid sequence of said *rhg1* allele or part thereof to a reference nucleic acid sequence.
12. The method of investigating an *rhg1* haplotype of a soybean plant according to claim 11, wherein said determining of said nucleic acid sequence of said *rhg1* allele or part thereof is a determination of a single nucleotide.
13. The method of investigating an *rhg1* haplotype of a soybean plant according to claim 11, wherein said determining of said nucleic acid sequence of said *rhg1* allele or part thereof is a determination of the nucleic acid sequence of an exon of *rhg1*.
14. The method of investigating an *rhg1* haplotype of a soybean plant according to claim 13, wherein said exon is exon 1 of *rhg1 v.1*.
15. The method of investigating an *rhg1* haplotype of a soybean plant according to claim 13, wherein said exon is exon 3 of *rhg1 v.1*.
16. The method of investigating an *rhg1* haplotype of a soybean plant according to claim 13, wherein said exon is exon 1 of *rhg1 v.2*.

17. The method of investigating an *rhg1* haplotype of a soybean plant according to claim 13, wherein said exon is exon 2 of *rhg1* v.2.
18. The method of investigating an *rhg1* haplotype of a soybean plant according to claim 13, wherein said determination of said nucleic acid sequence of said *rhg1* allele or part thereof is a determination of the nucleic acid sequence of a leucine rich repeat domain.
19. A method of introgressing SCN resistance or partial SCN resistance into a soybean plant comprising:
- performing marker assisted selection of said soybean plant with a nucleic acid marker, wherein said nucleic acid marker specifically hybridizes with a nucleic acid molecule having a first nucleic acid sequence that is physically linked to a second nucleic acid sequence that is located on linkage group G of soybean A3244, wherein said second nucleic acid sequence is within 500 kb of a third nucleic acid sequence which is capable of specifically hybridizing with the nucleic acid sequence of SEQ ID NO: 5, 6, complements thereof, or fragments thereof having at least 15 nucleotides; and,
- selecting said soybean plant based on said marker assisted selection.
20. The method of introgressing SCN resistance or partial SCN resistance into a soybean plant according to claim 19, wherein said introgression of said SCN resistance or partial resistance is carried out by backcrossing with a sensitive soybean recurrent parent.
21. The method of introgressing SCN resistance or partial SCN resistance into a soybean plant according to claim 19, wherein said introgression of said SCN resistance or partial resistance is carried out by backcrossing with an *Rhg4* SCN resistance soybean recurrent parent.
22. The method of introgressing SCN resistance or partial SCN resistance into a soybean plant according to claim 19, wherein said introgression of said SCN resistance or partial resistance is carried out by backcrossing with an elite soybean recurrent parent.
23. The method of introgressing SCN resistance or partial SCN resistance into a soybean plant according to claim 19, wherein said nucleic acid marker is selected from the group consisting of SEQ ID NOs: 54-1096 and complements thereof.

24. The method of introgressing SCN resistance or partial SCN resistance into a soybean plant according to claim 19, wherein said nucleic acid marker is capable of detecting a single nucleotide polymorphism selected from the group of those set forth in table 2.

25. The method of introgressing SCN resistance or partial SCN resistance into a soybean plant according to claim 19, wherein said nucleic acid marker is capable of detecting a SSR selected from the group consisting of SEQ ID NOs: 54-1096 and complements thereof.

26. The method of introgressing SCN resistance or partial SCN resistance into a soybean plant according to claim 19, wherein said physically linked second nucleic acid sequence is within 100 kb of a nucleic sequence which is capable of specifically hybridizing with the nucleic acid sequence of SEQ ID NO: 5, or 6, complements thereof, or fragments thereof having at least 15 nucleotides.

27. The method of introgressing SCN resistance or partial SCN resistance into a soybean plant according to claim 26, wherein said physically linked second nucleic acid sequence is within 50 kb of a nucleic sequence which is capable of specifically hybridizing with the nucleic acid sequence of SEQ ID NO: 5 or 6, complements thereof, or fragments thereof having at least 15 nucleotides.

28. The method of introgressing SCN resistance or partial SCN resistance into a soybean plant according to claim 27, wherein said physically linked second nucleic acid sequence is within 25 kb of a nucleic sequence which is capable of specifically hybridizing with the nucleic acid sequence of SEQ ID NO: 5 or 6, complements thereof, or fragments thereof having at least 15 nucleotides.

29. A method for the production of a soybean plant having an *Rhg4* SCN resistant allele comprising:

(A) crossing a first soybean plant having an *Rhg4* SCN resistant allele with a second soybean plant having an *Rhg4* SCN sensitive allele to produce a segregating population;

(B) screening the segregating population for a member having an *Rhg4* SCN resistant allele with a first nucleic acid molecule capable of specifically hybridizing to

linkage group A2, wherein said first nucleic acid molecule specifically hybridizes to a second nucleic acid molecule linked to said *Rhg4* SCN resistant allele; and,

(C) selecting said member for further crossing and selection.

30. The method for the production of a soybean plant according to claim 29, wherein said first nucleic acid molecule is capable of specifically hybridizing to said second nucleic acid molecule having the nucleic acid sequence of SEQ ID NO: 4, complements thereof, or fragments thereof having at least 15 nucleotides.

31. The method for the production of a soybean plant according to claim 29, wherein said first nucleic acid molecule is capable of specifically hybridizing to said second nucleic acid molecule having the nucleic acid sequence of SEQ ID NO: 7, complements thereof, or fragments thereof having at least 15 nucleotides.

32. The method for the production of a soybean plant according to claim 29, wherein said nucleic acid molecule is a nucleic acid marker capable of detecting *Rhg4* haplotype 3.

33. The method for the production of a soybean plant according to claim 29, wherein said first nucleic acid molecule is capable of specifically hybridizing to a nucleic acid molecule having a sequence that is present on linkage group A2 within 100kb of said *Rhg4* SCN sensitive allele.

34. The method for the production of a soybean plant according to claim 33, wherein said first nucleic acid molecule is capable of specifically hybridizing to a nucleic acid molecule having a sequence that is present on linkage group A2 and located within 50kb of said *Rhg4* SCN sensitive allele.

35. The method for the production of a soybean plant according to claim 34, wherein said first nucleic acid molecule is capable of specifically hybridizing to a nucleic acid molecule having a sequence that is present on linkage group A2 and located within 25kb of said *Rhg4* SCN sensitive allele.

36. The method for the production of a soybean plant according to claim 35, wherein said *Rhg4* SCN sensitive allele is also present in soybean line A3244.

37. The method for the production of a soybean plant according to claim 29, wherein said first soybean plant having an *Rhg4* SCN resistant allele is selected from the soybean maturity group consisting of 000, 00, 0, I, II, III, IV, and V.
38. The method for the production of a soybean plant according to claim 29, wherein said first soybean plant having an *Rhg4* SCN resistant allele is selected from the soybean maturity group consisting of VI, VII, VIII, IX, and X.
39. A method of investigating an *Rhg4* haplotype of a soybean plant comprising:
- (A) isolating nucleic acid molecules from said soybean plant;
  - (B) determining the nucleic acid sequence of an *Rhg4* allele or part thereof; and
  - (C) comparing the nucleic acid sequence of said *Rhg4* allele or part thereof to a reference nucleic acid sequence.
40. The method of investigating an *Rhg4* haplotype of a soybean plant according to claim 39, wherein said determining of said nucleic acid sequence of said *Rhg4* allele or part thereof is a determination of a single nucleotide.
41. The method of investigating an *Rhg4* haplotype of a soybean plant according to claim 39, wherein said determining of said nucleic acid sequence of said *Rhg4* allele or part thereof is a determination of the nucleic acid sequence of an exon of *Rhg4*.
42. The method of investigating an *Rhg4* haplotype of a soybean plant according to claim 41, wherein said exon is exon 1.
43. The method of investigating an *Rhg4* haplotype of a soybean plant according to claim 41, wherein said exon is exon 2.
44. The method of investigating an *Rhg4* haplotype of a soybean plant according to claim 41, wherein said determination of said nucleic acid sequence of said *Rhg4* allele or part thereof is a determination of the nucleic acid sequence of a leucine rich repeat domain.
45. A method of introgressing SCN resistance or partial SCN resistance into a soybean plant comprising:
- performing marker assisted selection of said soybean plant with a nucleic acid marker, wherein said nucleic acid marker specifically hybridizes with a nucleic acid molecule having a first nucleic acid sequence that is physically linked to a second nucleic

acid sequence that is located on linkage group A2 of soybean A3244, wherein said second nucleic acid sequence is within 500 kb of a third nucleic acid sequence which specifically hybridizes with the nucleic acid sequence of SEQ ID NO: 7, complements thereof, or fragments thereof having at least 15 nucleotides; and,

selecting said soybean plant based on said marker assisted selection.

46. The method of introgressing SCN resistance or partial SCN resistance into a soybean plant according to claim 45, wherein said introgression of said SCN resistance or partial resistance is carried out by backcrossing with a sensitive soybean recurrent parent.

47. The method of introgressing SCN resistance or partial SCN resistance into a soybean plant according to claim 45, wherein said introgression of said SCN resistance or partial resistance is carried out by backcrossing with an *Rhg1* SCN resistance soybean recurrent parent.

48. The method of introgressing SCN resistance or partial SCN resistance into a soybean plant according to claim 45, wherein said introgression of said SCN resistance or partial resistance is carried out by backcrossing with an elite soybean recurrent parent.

49. The method of introgressing SCN resistance or partial SCN resistance into a soybean plant according to claim 45, wherein said nucleic acid marker is selected from the group consisting of SEQ ID NOs: 54-1096 and complements thereof.

50. The method of introgressing SCN resistance or partial SCN resistance into a soybean plant according to claim 45, wherein said nucleic acid marker is capable of detecting a single nucleotide polymorphism selected from the group consisting of those set forth in table 4.

51. The method of introgressing SCN resistance or partial SCN resistance into a soybean plant according to claim 45, wherein said nucleic acid marker is capable of detecting a SSR selected from the group consisting of SEQ ID NOs: 54-1096 and complements thereof.

52. The method of introgressing SCN resistance or partial SCN resistance into a soybean plant according to claim 45, wherein said physically linked second nucleic acid sequence is within 100 kb of a nucleic sequence which specifically hybridizes with the nucleic acid sequence of SEQ ID NO: 7 or complement thereof.

53. The method of introgressing SCN resistance or partial SCN resistance into a soybean plant according to claim 45, wherein said physically linked second nucleic acid sequence is within 50 kb of a nucleic sequence which is capable of specifically hybridizing with the nucleic acid sequence of SEQ ID NO: 7, complements thereof, or fragments thereof having at least 15 nucleotides.

54. The method of introgressing SCN resistance or partial SCN resistance into a soybean plant according to claim 45, wherein said physically linked second nucleic acid sequence is within 25 kb of a nucleic sequence which is capable of specifically hybridizing with the nucleic acid sequence of SEQ ID NO: 7, complements thereof, or fragments thereof having at least 15 nucleotides.

55. A substantially purified nucleic acid molecule comprising a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 5, 6, 8-23, 28-43, complements thereof, and fragments of either.

56. A substantially purified first nucleic acid molecule with nucleic acid sequence which specifically hybridizes to a second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NOs: 5, 6, 8-23, 28-43.

57. A substantially purified nucleic acid molecule comprising a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 7, 44-47, and 50-53, complements thereof, and fragments of either.

58. A substantially purified first nucleic acid molecule with nucleic acid sequence which specifically hybridizes to a second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NOs: 50-53.

59. A substantially purified protein or fragment thereof comprising an amino acid sequence selected from the group consisting of SEQ ID NOs: 1097, 1098, and 1100-1115 and fragments thereof.

60. A substantially purified protein or fragment thereof comprising an amino acid sequence selected from the group consisting of SEQ ID NOs 1099, and 1116-1119 and fragments thereof.

61. A transformed plant having a nucleic acid molecule which comprises: (A) an exogenous promoter region which functions in a plant cell to cause the production of a mRNA molecule; (B) a structural nucleic acid molecule encoding a protein or fragment thereof comprising an amino acid sequence selected from the group consisting of SEQ ID NOs: 1097, 1100, 1098, 1101, 1102-1115; and (C) a 3' non-translated sequence that functions in the plant cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule.

62. The transformed plant according to claim 61, wherein said plant is soybean.

63. The transformed plant according to claim 62, wherein said plant is soybean is selected from the group consisting of PI548402 (Peking), PI200499, A2869, Jack, A2069, PI209332 (No:4), PI404166 (Krasnoarmejkaja), PI404198 (Sun huan do), PI437654 (Er-hej-jan), PI438489 (Chiquita), PI507354 (Tokei 421), PI548655 (Forrest), PI548988 (Pickett), PI84751, PI437654, PI40792, Pyramid, Nathan, AG2201, A3469, AG3901, A3904, AG4301, AG4401, AG4501, AG4601, PION9492, PI88788, Dyer, Custer, Manokin, and Doles.

64. The transformed plant according to claim 61, wherein said promoter is an *rhg1* promoter.

65. The transformed plant according to claim 61, wherein said promoter is an *Rhg4* promoter.

66. A transformed plant having a nucleic acid molecule which comprises: (A) an exogenous promoter region which functions in a plant cell to cause the production of a mRNA molecule; (B) a structural nucleic acid molecule encoding a protein or fragment thereof comprising an amino acid sequence selected from the group consisting of SEQ ID NOs: 1099, 1116-1119; and (C) a 3' non-translated sequence that functions in the plant cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule.

67. The transformed plant according to claim 66, wherein said plant is soybean.

68. The transformed plant according to claim 67, wherein said plant is soybean is selected from the group consisting of PI548402 (Peking), PI437654 (Er-hej-jan), PI438489 (Chiquita), PI507354 (Tokei 421), PI548655 (Forrest), PI548988 (Pickett),



PI88788, PI404198 (Sun Huan Do), PI404166 (Krasnoarmejskaja), Hartwig, Manokin, Doles, Dyer, and Custer.

69. The transformed plant according to claim 66, wherein said promoter is an *rhg1* promoter.

70. The transformed plant according to claim 66, wherein said promoter is an *Rhg4* promoter.

71. A transgenic seed having a nucleic acid molecule which comprises: (A) an exogenous promoter region which functions to cause the production of a mRNA molecule; (B) a structural nucleic acid molecule encoding a protein or fragment thereof comprising an amino acid sequence selected from the group consisting of SEQ ID NOs: 1097, 1100, 1098, 1101, 1102-1115; and (C) a 3' non-translated sequence that functions to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule.

72. A transgenic seed having a nucleic acid molecule which comprises: (A) an exogenous promoter region which functions to cause the production of a mRNA molecule; (B) a structural nucleic acid molecule encoding a protein or fragment thereof comprising an amino acid sequence selected from the group consisting of SEQ ID NOs: 1099, 1116-1119; and (C) a 3' non-translated sequence that functions to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule.

**rhg1 LRR**

consensus LRR	L F S N L P	N L E E L D L S N N	L T	S L P P G
	a x x x x	x L x x L x L x x N	L	I
LRR 177-200	T L G L L P	G L R K L S L H D N Q	I G	G S I P S
LRR 201-224	S L G F C P	N L R G V Q L F N N R	L T	G S I P L
LRR 225-248	L A N S T	L L Q S L D L S N N L	L T	G A I P Y S
LRR 249-272	L T H S F	K L Y W L N L S F N S	F S	G P L P A S
LRR 273-297	W G G N S K	S L T F L S L Q N N N	L S	G S L P N S
LRR 302-325	L G S L R	R L Q N L I L D H N F	F T	G D V P A S
LRR 326-349	I G T L S	E L N E L S L S H N K	F S	G A I P N E
LRR 350-373	T L S N L S	R L K T L D I S N N A	L N	G N L P A
LRR 374-397	S L G R L R	S L T L L N A E N N L	L D	N Q I P Q
LRR 398-421	S I A N I S	N L S V L I L S R N Q	F S	G H I P S
LRR 422-445	S F D S Q R	S L R Q L D L S L N N	F S	G E I P V
LRR 446-470	L L A K K F N S L N	L F N V S N S	L S	G S V P P

**Figure 1****Rhg4 LRR**

consensus LRR	L F S N L P	N L E E L D L S N N	L T	S L P P G
LRR 34-57		H V T S I S L A S H S	L T	G T L P S D
LRR 58-80	L N S L S	Q L R T L S L Q D N S	L	G T L P S
LRR 81-104	L S N L S	F L Q T V Y L N R N N	F S	S V P T
LRR 105-130	A F A S L T	S L Q T L S L G S N P	A L Q P W S F P T D	
LRR 131-154	L T S S S	N L I D L D L A T V S	L T	G P L P D I
LRR 155-178	F D K F P	S L Q H L R L S Y N N	L T	G N L P S S
LRR 179-203	F S A A N	N L E T L W L N N Q A A G L S	G T L L	
LRR 227-250	L S N M S	A L S D L Q L R D N Q	L T	G V V P A
LRR 251-274	S L T S L P	S L K K V S L D N N E	L Q	G P V P V
LRR 333-356	F G K G V N	K I I T V N F E K Q G	L Q	G T I S P
LRR 357-380	A F A N L T	D L R T L F L N G N N	L I	G S I P D
LRR 381-404	S L I T L P	Q L Q T L D V S D N N	L S	G L V P K
	F P P K V K			

**Figure 2**

<110> Parnell, Laurence D.  
Hauge, Brian M.  
Parsons, Jeremy D.  
Wang, Ming Li

<120> Nucleic Acid Molecules And Other Molecules Associated With  
Soybean Cyst Nematode Resistance

<130> 38-10(15810)B

<150> US 60/174,880

<151> 2000-01-07

<160> 1123

<210> 1

<211> 127197

<212> DNA

<213> Glycine max

<223> Seq ID: 515002\_region\_G2

<400> 1

agagcttgaac agtatatgga ttagccacca tgttgaaagg cacgtaaggc caaatctcag 60

cctttttccc tgttgacttg gccttcttta agaccttggt tggttccaca taaccagtta 120

ctgttacttt ctgctgcttc ctgtttatct ccacggatcc tactcctaca ettacaccac 180

catttttttag ttaatttcat caccttcata cgtgttgaat gctatataat taatatatca 240

cactattaat ttccacaaa agaaaaatat tacactatta attgataaca tactctctaa 300

cacttttggt ttaacatat attttattatt aattaaaatt tattgaaaac cataaattat 360

aaattataag tgaaaccgta aaataaagag tcataatcaa cattttttgt gattttcaac 420

aaattttaac ataaagatg gttgtgctag ctttttttat atatattatt ataattataa 480

tatggcacac taccatgga agctgatttt gagaataaaa ttgagggaaa acgtagcact 540

taccatctaa tgaagagaga gttttcctaa ccttgagcac acagccgtca caatccatca 600

tcaccttcag ctccacggtc tgtaattgct tcttgttctt gttcttgggt ttgtgttggg 660

gggtgccatt cccacttccc attaaatcag accagtaato tccaaccccc atttttgtgt 720

atttaatatc aaaaccaata tgaaagcaag ctgggaagtt cttgatgagt agagaaaatt 780

agtaggtatt agatgggttg aacatcattg atgggtgtgg agcgtaggag aagaatttat 840

atagagaaaa agcaagggtcc aaaccacaa gaagttacaa ggaactttct tgcaaaacag 900

aaaaatatc ctccactcact taccttctaa tgatctaaaa accaatgctg ctctttgaag 960

aacttttttt tagttacctt gggatatttt taccacatct aactaaaatt gattttggta 1020

gaagtaaaaa ttaatcttat ttgttttaaat tttatcataa ttttaaaaaa aatttaaacg	1080
tacaaaattt agttaaactc aagattaatt cacatcagca tagtctacca gaattttgaa	1140
agttattcac acaattatat ataggctttt tcaccattca gattcaatga tatgaatgga	1200
aaggatatgt ggtccctaac tacgccagtg aataaagaac cttagaacgg attataactt	1260
cttaacggag aaaattttta cgtgggggttc aagaaagtag ataagaagga acgcattcct	1320
gtatcacctt ttctattcgc aaatatatat gtaaatagta aaaattttat gtactgcttg	1380
cggttcagtt gacacttgac aaagttattt atttgatatg taaaagaaa tggacatatc	1440
aaatgcacgc taatactgag aactctctta ttgaagaca ataattttat gaggacgaaa	1500
gtgaagact gaaagcagct tcaccgact catgcacata tcgatcgaa tgaacaacat	1560
aatgtggaac agataattgac ggggattgta ttggaagta tgcaagtgca agcatttagc	1620
accattgaca aagataatcc ttcatattgc aacggctatg agccttttac cctctccatt	1680
gataaaattt cagtcaactt aaggccaatc aaactcaca atatattgtca agtttgact	1740
ttgcaaaaaa ataataagga aattatctat caaactaaa taaatttcac aaaatcatca	1800
ttccttagcc agactcatga tccaaactca tatttttaat taaatttcag tctcttaaca	1860
ttaaggagtt ttccattttt tcttgatttt aaaaagaaaa tatttcaatc gtctctctta	1920
gcatgacctt ttggactcgt tgtgacaata ttcataaac tttctcttaa aatctttcga	1980
cctttcacta taaatacttg ttaattgtat ctaaaagtat gaagcattat ttgttaagca	2040
aattctttta gcttagtcac gaagaagtta ttggaagtg ctaataattt tcatcatgc	2100
atggttcaag aattccatta gacaaagaaa attgtctcag tagttatcat cgataacaca	2160
agctactaga ccaattgacc aaagaagttt cgtcttttaa attcttttcc catcttgttt	2220
gaatatgaga tatttgattc tcattgatac tttttcgata agataaagg gacatttttg	2280
cttctgtttc taagatattt ctataaaggg aaccaaaga agcaatttga atcattaatt	2340
taactaaccc actaataat ctttttttca attagcaaaa ctgctgaaag ctgagttcca	2400
cttttggaac acaaaagtag agtgataagg aaaagataaa aataagaaa tgcaacgaaa	2460
agtggtgggt agaagacaaa agagaagcat ggtcaggat accaccctaa actaaagtat	2520
gtgcttaagt ttggaagtc cggaggata caccaaaggc ttagtttcac ttgcattaac	2580
atattaataa taattcacaa gagaagaaat ataaaataag ataaatcaat tttttcaca	2640
gttaaatatt attttatttt ttggaatttt tacaagaaat aaaaaaatc ataacttaa	2700
attaactcat gtatgataac tttttaatag ttttttggtt taattttcct aaaagataag	2760

gcgtataaat taatttttagt ttacgagaga agttggattt gtttatttct tatattaagt	2820
gccttgttta gaaattttatt gaaagtggacc caaaatcaat tatatgtgga ttttaaacact	2880
tcaaagctcg ctgaaactga aagcatttca ttcaaaagta aatttaacgt taatgttccc	2940
ctgattagtt gagaaaaagga catgaattgt actagtatta atttacctcg agatcaaaga	3000
gagcaaaaaa tttaatcata ctatatttac gcagaaaaag atgataaaca agtgatatagt	3060
acagacgtga gacgtggaat aaaactttga ttagtataaa attgggtcatt gttgctttta	3120
gaaacttgct gtggacacac atgacaaaaa ggatttgcatt ttttatgaag ctogtctcat	3180
gtttaaactt taatttggac ggaattgaaa aacatctcat tgatatattc atgtgctttc	3240
ttctgtctag cttttggatt gtagatactc tttgattatc ataaaaggaa cgaacgactt	3300
taatttggaa attgagaccg aagttgcaat agaaatcgtt acacgtcttc tctagtgtgt	3360
atgtatttgt tccttttggg tttaaaggat gaccaaagt tgaggccagt cacaataaa	3420
agggaaactca cgaagctggg gaaagtgttt agtatagttt aatgtttaaa ctcatgcacg	3480
caccagatat cttaaagaatc tataccagat atataacggt tgtctatata gtttctttcc	3540
attcccattg ttggagaatt aaagcttctg attatttagc atattttatt gtataatttt	3600
actatatagt attttgtact ttaagtatct tatcaaatct tcggttgcaa catcattgga	3660
gaagatatag tggcattcat aatgaataga agtccaatga attcacttta tgttgtcaac	3720
tagtagtcca atgaattatg ttgtggttgg cttaagaatg aatacaagtc aaatttcacg	3780
tggcttacag tcacagtggc atggaatata taacagagtg ttgcgtcaaa aaacaataa	3840
tgacgacaac aggactaaat acaatgcatg tatttttttt ttaattgaga aataatttat	3900
tctttaaagt tctctaaaaa aagtaacatt ttactataaa ttttccccac tcgaattgag	3960
ggtattaaag taattttcat tgtttttgct tttatttttt atttaaaaaa catttgttca	4020
gttaatgaga ttaatgaatt ttgtcacttt tgggtaacac atgtaaaaat gaagacggat	4080
gcagaaacaa acacctacta aatgacagta tatgaaacct ttatatcaag tacatatatt	4140
attttctcaa ttttcttttaa ctttaactca ttaagccaat cgatctatta catctottac	4200
tttgatttga tgtttcactt atcattaatt aaagataaac aatgtattac attttcttaa	4260
ctaaaaatta gcattctaaga catatatatt tgtagttcat ttgtcatga aaaaatggg	4320
tacatcatcc gatacaatta ctaaaactaaa ctaagggaag aattattctc acttttatat	4380
ttgttattta aactatcaat tctgaaattt ctattttgcc cactaaccaa attcctccac	4440
ccctctctt ttcctctccc tccttctcat atccaaaacc gttgccctca caccaaccta	4500

gactgcatca	ccccccctc	agcgccaccc	ttctcccac	acaaaacgtc	gtcaatccac	4560
gaccacgcc	tcctccctc	cgcaacccat	tgtcaatcca	catgatgcc	atctccccc	4620
tccatgcaat	tcgttgatcc	acacagcacc	ttgcccac	aagatcaagc	acccttggtt	4680
ctcttttttg	gattttattg	caggttttgt	tgtgttgtaa	tcgtatttct	gttgcatcat	4740
tttttacacg	ctactgtctg	taatggaaat	agaatgaaat	cgtatttctt	ttgctttttt	4800
ttctttacc	catgtgtgca	acgaaaatag	aattttgtta	gagacttatt	taacggagtc	4860
atattttcgt	gggaaggaca	agctaataaa	aaaatatattg	aaactaagag	ggattgtgcc	4920
aatagcaata	ttggattatg	tagttccctt	gaactacata	aaaccttctt	cgtattgaaa	4980
tatgcctaat	taaaactttc	atacctgaaa	tatgggtgtc	ttcaatttat	tagctaatat	5040
ttattttttt	taatacaagta	gttgaaaaaa	aaattcagtt	tcattttatt	atttatattt	5100
ttttcatcaa	gtggtaatat	gacatacgga	gtgttgccct	aacctgtctt	gttatggata	5160
actcatactc	atgtttatcat	cttcgatgat	tgtgacaata	acatgttggt	aaattgaaac	5220
ataaaaaaat	tttagataac	tatttgacaa	aaaatgaatt	tttaggtagt	aatttgaaaa	5280
taattttatt	tccaggtagt	aaaaacttaa	ttaaatcaat	ctgaaaaata	attttttagc	5340
taacaaatga	taaggagaaa	ttttataaaa	ataataattt	taatttcaaa	taaaatgata	5400
aattttatta	tttcattaaa	aataaaaaa	aaaatggtaa	tcataaattt	aagttatggt	5460
taaaagttaa	tcttataagt	caataagaaa	aaattatttc	tcaaacactt	ttatttgatc	5520
aaatatttgt	aagtttgtgt	aaaaaactaa	aaattaatta	aaataacctg	atgagcatat	5580
atgtaattta	cttttatata	gacttaaaaa	actttatctt	tttttaagat	aattttctcat	5640
tcaaaaaaag	aaatacatta	aattaatgag	atattttttt	attttagtaa	ttaaaaaaag	5700
tactatttga	aattcacaat	aacacttatc	aatttatcat	atttaacag	ctgaactaaa	5760
ttataaaaa	aattattaag	attaaattaa	aaagttgagt	tacatacatt	gaaccaacaa	5820
gatcaaacac	tgagatatat	attttggatc	cagagaataa	taaagggaag	gtgaagggaag	5880
gcataacatg	ttacatacat	tgaaccaaca	agatcaaaac	gtgagatatt	aattatttat	5940
tgtttaata	aatttttaat	acctgaaaaa	catatcattt	ttaaattatt	acttaaaaaa	6000
tcattttttt	caaataccta	ctagaaaaaa	aaatttagtt	ttattttatt	atgtcattaa	6060
tcataatcac	gtccactcgg	tcatatgtca	ttgaaagtga	taacgtaaca	ataaaatatt	6120
cgtgacaagg	tgtgatgacg	taacattttg	tatgtcacat	tatcatttca	ctaaagacaa	6180
ctaacaacaa	ataataaaat	aaaattgaat	tttttttaaa	atacttaact	ttaaaaaata	6240

aataattaaat aataatctaa aatgatctg tatgtccatt atcaaaaatt taagtaagcc	6300
gaataataac cttcttatc agtcatttg gattcgctaa caactcgtc tgatcgagcc	6360
tatagtaatt agctctcta gctagaaca ttgcccctg attatcaata tgatgttctt	6420
gcacaaaacg acaaaatgat ttcattgatt ttatgataat aagcgcagga ttctctatta	6480
ttcagagcaa ttaaatgaac gcaattacga ggcagcaaat gcaaaatgct tcacatgagg	6540
ctattggcta tttaaaattt cactttgaca tttttataat taattttgtc tcaataattc	6600
gcatgacaat tagcattaaa ggtagctaa aactctaatt taattttgta caatatatta	6660
aaaaaaactg agaacaatga ttttaagtttt gttattatat caaatactac tataaaatag	6720
aaaaaactta cacccttcaa atgtactaga ttccggatttt ttttatagaa aatattttta	6780
aattaaatgt cataactcaac agtcaacacg gtatttcaca gtcccctaga atttcaaca	6840
aatcaacca agtaacataa attagttgac tgaaaaaatg aattaaatta aaaggcagtc	6900
atgggataac accaggcata ttaatttaca taacctacc cgagcaaaag cgggtttaac	6960
agataatggt ccagtcagtg tttaggaatct aacatgctgg caatgtcaaa aaaataacag	7020
cattggatgg tgtaagatct aaaatactta caaagctagg aggaggacaa aatggataca	7080
tatttggatg tacatgtaat aactctatct agacaggcta gttgagatac ataagaataa	7140
gaacgtgtct gtctcagtaa agggcagaca caagtagaag tagaagaaac aaatagcagt	7200
gccaatgtac ccggcagcat gaaatcatcc gagatggagc agccgaaggt ttgtggggag	7260
ctcttcacgc aacagctgga gcaactgcat gcccgttcgt tctttgttgc tcactgttag	7320
gcaatgggtt tgatgaagtc tcatttggag aaaggggcatt cctcttgctt agactcagat	7380
cctgacctac agatgcatgt agactgtata taagcaaaag gaataaaaag ggagacggga	7440
agaacaggtg taaggtagaa aaaagccttt gcatcaagca ccaggcaaat ggttaagaga	7500
ccaagaactc acaagaagtc agcttcattg cctaagtaga atgattagaa ctaaagctaa	7560
aatatattag cttataaact caaagtacta tgactcaciaa tttagcgtg accacgctag	7620
cttcttgttt cccctatcaa ataccaaacg gtatcctgtc atgaagtctt ctgccaaaa	7680
aattttattg ttttaagatc aaagtatctt ttaataccat attccagagt atgggtaatc	7740
agtagacttt gataaggaaa atatttaact tacgtccgat tgttcccata tctccttcag	7800
ttggctgtat ggctaaacaa aatccaatga ctccctgcac aggaaggacc gatctaacta	7860
atttagctac aagccgacaa ctattcttat aatgaagtct ctgttaacat ctttaattat	7920
atgtcaaaat tttagtccag gatcacttaa ttcacacca cctcattgcc atagaataca	7980

aatacggggt catagaccac aaaactgtta ttctgttga acgtgagagt taagctggga	8040
accttttgca actcttgcga actgtctcag aaaagaataa taaataata aagccatcaa	8100
agagaccaga aaattctacc aattaggaaa tcatgcacca acgcaagagg gaagagagac	8160
agagatctat ccagaaaact cacttgggaa cataacaata ctcccaagga gatccttcaa	8220
agctagatct tgaaccattt acttgttgat caaactatca caaataaaa catattaaaa	8280
gactgagcag aaatttcat taatggaatt tagaaaacta aagtacctcc tcagctattg	8340
ctccatacac atgaccagga agaaaagtaa atgatgtccc gctatcaacc tgcactttaa	8400
aacttgtcat ttttaagcaa gaattcccaa cacacatga ctccactcca ataatgtagg	8460
ttgaactgac caggaaattt atcataatta gcactgtgac accattttct aaagtataat	8520
ttagtgcagt gaaacattgt aaactaatat aaagtacaga atttcatact ataatccatc	8580
caaaggcaag aatgaagtag actgttggat ggttggcccc tggccccaa aaaatattct	8640
accagaatca tcttcattaa agcacaagga aaaagaatcg tggattaatc ctgatttagc	8700
aagaaaactt ggaactgaac tctccccagg ccccaaacct aataggccat ctggagcaac	8760
cccatccaaa taaccaccac ttgtcttcac accacacctg cattcaacaa aactgacgtc	8820
agagaaaaaa ccaggaaact tgtttttaat taaaaaaga atacatcatc agggaaagt	8880
atgttgaaat aacaaagaca gggcagtcag ccacacccaa gaacaactgg agcctgaaca	8940
gatgagtttg ataaactgcc gcctgactga agatgcagta tgtcctcaac caacaatcca	9000
gaactcgacg tgttctcoga caagtaactg accatgtatg gacactgctg ctgtgaactt	9060
ttacaattcg aaccttatac acacaactga tgactgcaag atagatgctt actggataag	9120
gaccgagacg gactatactc attcagatct ctatcctgca taatagatgc cccatattgt	9180
aaggagcacc tgccaatcaa ttactggaaa taagtgtctaa acctgttata gattctaacc	9240
ctcattgaaac cattacactc agaaaggaac catgttttgt gttagtgtca cattaaactc	9300
gaattgaaac cttatccaat caaggatttg caattcgcat gattaactat ttgttaaaaa	9360
atcaataaaa caagctaata taatccgata ttttattatt ttattacat ttaagatatt	9420
gagactacaa gttacatagt agagtaaacc aacatttttag ttctggaag tataaagcct	9480
agtcacataa acattagtcc ccaaaactaa gaaacttcaa aaagtcctc gaagctgcaa	9540
tcgcgaatc gcattaatcc aaagttataa aaaaatatgt gacttaatga taatattatc	9600
atatgtttaa gagaccaaaa tctcaagatg aaacaactag aaactccctg gactaatttt	9660
aaattttcct tagtttgaag aactaatgtg acacctcggt atgcttaatc atagtttact	9720



ctacatagta agagaaatca aagaaaaaaa tagattaggt atgatattca taccagattg	9780
gaatagtaac tcgatgacaa tggagcacac tgtacgcaat cacacggaat ccaaaggaga	9840
tcactccctg cgtccaacgc cagagaaaac gaagtgtctg gtgttctcat atcaatccac	9900
gtgtaatgca acctaatac cagagacgca ttatcagatt caaacggaag aaaaggagc	9960
aattagggat tattattatt attattatta ttattaccag ccgaagtcgt tgcgagcgga	10020
catcgttttg ctaccgtgag aaggaaacag tagctggtag cgtgcgcctc cgactttgat	10080
cttgcgcctg agaatgtgc cggtgagaag catcggtag tagcccatgc tccaccggtc	10140
cggccagtat ccggtcggag gccgaacagg ttctatttgc tcggcggaag ggtgaacgag	10200
gcgggcggaa aacgtgatcg gaaccggcat tgcgcgagcc gttacaaaa gcaacaaaag	10260
cagcaaccgc caccgcacgt agatcgagat ctggcacttg cacttattct gatgcctcgt	10320
tttaactgat ttaagtaacg attagtgtta attagtggag tgagggtgcg cagtggtcat	10380
catcatcgcc atggatcgta tcgtttcgtc cctgtgtggc tgtgtgtgag tgagagtggag	10440
agtgagagtg aggggtggata aaacaaacaa acaaaactag cgcatttttg tgcgggtgga	10500
attagactgt tactaagtcg ttaattaatg gggaaggaa agtggtaga ttagtgtttg	10560
taacagtaag tgattattgt aaatgatgat taggaggaat aaggggtcaa cactgcagcg	10620
acgaagcgaa acgtcacgcg cggtygcccc accatgtctt tacgtgcttg agaataaac	10680
ggccttttat tgcgatgtc gatttgtctt tgccactgtg ggccecccca cttttattat	10740
tattctcttc cttttacgaa ataaaaata aaaaatacaa caacaaggc aaaagggttc	10800
ttaagtattt agtttcatta tataaataaa ataatagcct agatctagta aataatcaca	10860
ttatgtggtg tgggtcagga ataaagcttc acacacgaaa aaagaaatct tgcaagtaaa	10920
cagctgaaca cattaatgtt ttttaagaaa atctaaagtt attgaagaaa acaactgaga	10980
catgataatt tgactaatta atacttttag tgaaggagac gtattttaaa agataaagta	11040
taattataat aataattaat aaaataata acgattaata tttagtaatt tcattctatg	11100
taatatagtg atgatctcaa ctcaactgat aattttcaag ataatagtta taattgcact	11160
ctgtggaatc ttaagttctt tctccaaga aaaaaaaaa cttttttctt tcccctgtc	11220
gtgttctctt attctgccat ctccaattct gttcacaatc gtagggtgtg ccgccaatga	11280
gttttaatga taaagatcaa atacgtttgc aatgaatcgg gatgacaaga ctgagacaac	11340
caatagtgga agctaaccaa tgcacaagtg ctccaatcaa taaacaggc ccaaaaagggt	11400
ggggtgtgcc aaaatgtgaa ggtaagttta agtaggggtg tcacgccttg gattgcgtct	11460

gtgtaaatcc gtcaccaat ccaacaaaa aatattggat ggattttgtt gtttttcttt 11520  
 ttaaatcgac ctaatctgat catgaatgaa ttgatcggag atggatttgt tattaataaaa 11580  
 agttcaaaaa taattttctt aaatttttta aaattttttt tagaatttca aatacaatta 11640  
 ctgttaatat agttgcataa aaaaaattaa ccaccaattt caatgcacat attaactgca 11700  
 tcataaaatc aaattgaaaa caagtaacca acaaacattt aatttataaa gcaataata 11760  
 ctaaatcaaa tttaaccat aaagcagata acaaatgttc ttgaaaactt agtaacttta 11820  
 taaagtacac actagtacaa aataaactta aaatcatccc aaaaaatata taatactaca 11880  
 atagaaacac tgcaatatag tgataatgtc agacaattgc tcaaccagcc aacctcacac 11940  
 atagaaacac ggtaagcaaa agatcaaaat caatttattt actaataata aatttaaatt 12000  
 atgctatgca gaaaaagaa atatgcaaaa aaagaaatca tatcataaac taagttaaaa 12060  
 atattacctt aagaactaat agtctaact cccaatacta atactcctaa gaatagtcca 12120  
 agtagtaate ctaacactaa cattatttaa agtcaaacca tacaacttta aaaaatgttt 12180  
 taaaagtttc atcataacat aatatcaatt tatattcata ttgtaaacaa acggaaaaaa 12240  
 aaaaagaac tattattgaa tacctagttc catctttttt gtttcactca attcaactcg 12300  
 taaatcacgg acatttttgt tattagtgtt gagtcaattt tgggtacaaa tcaaaagttc 12360  
 aacagtaatg ggacttaaa aactacaaaa atggatcaag cactcaacct ttgtactaa 12420  
 atgcagactc aaatgacaca atagacataa gaatgaccaa tatatctcta gccatgaag 12480  
 aaataacatg atattttgat gctttctatt tccacatgc caaatgtca aatccaagac 12540  
 cgtcatcttc attgtcatcc tttaaatata tatccaactc actcctttgc tattcaccac 12600  
 attttttatt cattttcaat ctaaaatggg cgtccaactc ctcactctca tcaacatcgt 12660  
 tggcattacc ttgtgaagca tggatgaag ccaagttact agaattacta ctatcaatgg 12720  
 aataggatg ttctgaagca tattcaacaa acattttttt tataagatca tccaattttt 12780  
 tcagcatctc ttgggtttgg tcaacaccat gcattttctt aaaacaaaac tcaatataat 12840  
 caaatttata acacagatca agaaaagcag tcacaaataa aagatagcta atctgatcac 12900  
 tctctcaata cttgttaaac ttgagttgca tattagtgtt ctcttttgta tcacggatc 12960  
 atccatcatg ctcacatctt ttaggcattt ctgaatagta accaacttct taagaaatt 13020  
 cttagctgta acatgtagtg accagaaaa aaaaattgca tcatagaaaa ctttcaaaaa 13080  
 actcacaac acacagcatg gtttctaact catctcttta ggacatcttc cttcactatt 13140  
 tagaagatg agcacatatg cagcctcaac atactcataa cgattgaag cttgttcaaa 13200

ttttcagcaa catctaacat caaataagtg gagtgacca gaaaaaaaaa ttgcatcata	13260
gaaaactttc aaaaaactca caaacacacg agcacggttc taatccatct ctttaggaca	13320
tcctccttca ctatttagaa gagtgagcac atatgcagcc tcaacatact cataacgatt	13380
gaaagcttgt tcaaattttc agcaacatct aacatcaaat aagtgagggt ccatctgggt	13440
ggcacattaa gtgttagcat tgcctttgaa ttactactaa cgtcctccgc acacctcttt	13500
aaagtagtca acctacttgg agaaaacctc acaaacctac atgcaacctt tattttacta	13560
attgaagagt caatttcctt aaacctcatc catacaatca aattcaaat gtgtgttggt	13620
aactcattgg caagtggaact aaattgtcac aagtaataaa ataaaataaa agtctgagta	13680
tcaagtcac aggagcttta ttgtactta aattaatgta aacccaaatt ttcaacaag	13740
agataaagaa ttgagagtga aaagttaag ggctagtaaa aggcaagaag ataagaaat	13800
aaacaataat ttgaatgcaa aacatgactt caaaatgcaa ttgtgttgaa acttagcatg	13860
ccaaaactac ttgtgttaca atgttaatat tttttctcta tctaataa tccccgtgt	13920
tactcgcat tactatgata ttctatcttt ggaacccgc atcaagagcc taatttatct	13980
attttttctc ccaaatctct ttgcaagat aaaatagtaa atcatattaa gattaagat	14040
gtatgaata agctaaacaa acatcattct attcctagac atgatttcat ttagatgttc	14100
tttcccaatt ctctaggaaa caacctttt caaatgcatt atctcctaaa caatacatga	14160
gcatgggtga ttaaacaca atcaataata ataaagcaca taaaagaaca atagaactga	14220
aattgcatta aataaatagt aaggaagagt tacattacaa gagtttgggt tatcaggctc	14280
tcaacaatga ggggtttagc ctctaccatg agagacttta cactttaagg gtgtatggag	14340
gtaaagaag aaaagggtg aataaagaaa gagaaggaga gaatggttat agaaggagg	14400
tttccctat tagagatgtt caggctttgg gtgtattcaa gagagagttc taagtttcgg	14460
tgtgtcttct tccttttgc ttgaagtgc taaaattcac acgatctcac gctaagagcg	14520
cccttctggg cttagcaggt atggcggtga tcacgcacta agtcgcagcg catctgggct	14580
tagcgagtat ggtagcaagc acgcacttag cgcgagattc acactaaacg cgctttggg	14640
cttctcatg aatcttcttc acgctaagct tgatctgact actaagtga acgacgcgct	14700
aagcctttct tgtatgctaa acaagctgtc tccatcttca attttttttt aagacttttt	14760
cttcacgttt ttgcatcaat tttttctccg gcactgttaa ttctctctt ttgaatcttg	14820
ctgatcaaga atcaaaatga tatataaat ctcattattt cattgaaaa aacagtaag	14880
tagaggaatt cttagtcaaa attgactatc aattaaacc aaatttccca gttatcatgt	14940

gctgaacgct gcatatgtat aaactcttca ttcaacaaag tatgatcatt tcaaacacta 15000  
 agacctctac ttaaataatga aatggctaca ctgtagcac ttgcattatc aactgtttata 15060  
 caacaaattt ttgaatccc tcattctttc aggcaattct ctagggtctat ccaagtgggt 15120  
 tcacctttgt ggtcaccaat caaacaaaat tttagtattt tcttatgcaa cctaaatact 15180  
 tcatcaatgt aataagttgt cacacacatg taattcacat ttggattaa tgcctcatgta 15240  
 tcaattgtaa gtgaaccat tgtttatttg caacaatat gcattccaac tgccttttcat 15300  
 tattaacac ttgaatgcaa tatctagcaa tagtcacatg agaagaata ttaaaaaaag 15360  
 gttgagccac ttccataaac ttccataacc tttcatattc aacaaattta aaagaaagtt 15420  
 catttgtaat gatcatcttt gcaagttcta taatagtctt ctcttgattg aaatcaacaa 15480  
 atttcacact aagagcatta gggtcaccc cactttctt tccaacata attgtttttt 15540  
 tattttctat caacctgcct aaagggtttt tcatacatga agctaaatgt ttgttcaagt 15600  
 ttttggttcc atatttagaa ctattagcct tatagcattt ttacaatac ttgcaaaagg 15660  
 catcattatt taccctatca aaatgattcc acacttcaga aggggatcaa atagagcttg 15720  
 cgacccttgt tgtgtgggaa cctccagcac ttgttttgtt ttgattgtga tggagcacca 15780  
 gctgctacag gttcagtagg ggacattgag agggagtaaa cgggtggcaa ggacttaaga 15840  
 cattgacact tccaataaaa ctctattttt ccatttgtca aatttaataa ttaacaacc 15900  
 gtcaactaca agtctacaac gtgtatcata taaaaataa tgaataataa tataaataca 15960  
 actatataat ttttcttaag tgaataacat ctacaataa cactcaaatg agtcacttaa 16020  
 gcataattat taatacaact atatattttc ttatgtggca agaacattat actttgttta 16080  
 gagttgtaca aatacttact tagtagtaga caagtacta ggaattttat tacciaaac 16140  
 ttattccaca aatccaaatt ccaattctat gttctaacct ctacaacta tatatatatt 16200  
 ccttgtcatt tgatgagat gctttgattt cctatggaat ctcatctttg tctctaattt 16260  
 gtattcattg aaactcttga attagaaaag ccacatgtac ctagtatcat ttgatcataa 16320  
 ttcatatcct tatgtgatcc aaattttctt tgctaataat gttatttaaa acatgtacta 16380  
 tctctttatt tctacatcaa ccaagaatgt atttcttgat attcatacca gatttttgtc 16440  
 acactgttcc ccacagtgat agttcctaaa tataattttt tatgagtcga agacaataa 16500  
 aaaaccacat ttataacaag caatacaaac aaagcagta caatatagca atctatcata 16560  
 tgctgtaaga gaataacagg aaggtaatca taaaatgaga taaaggctga ataatttgta 16620  
 aaatatctcc acctagagag tataagtaca gattttaata accacatcaa aagaataaaa 16680

aaaaaattat gtttcagaca cattcaccac agtgatgggt tcttaaaaaa aatcatctat 16740  
 cgattcatat gtaccaatta ttaatacaat aatgcagtta aatatgcagt aagcattttg 16800  
 tctaccagtg ccatgtagac gtaggagagt gacagagagg ttgcggcagc atgcgcggcg 16860  
 tagcgcagtc gcgcagagag attgagagag gatggagatg cgaagcgcca aagagaaagg 16920  
 cggagcagag acctgagaaa ggacgcagtc acgcagtcgc acaaccgcga gggaggaaga 16980  
 ggcggagcga gcgcaacggc gcggaggagc ggagatgcga agcatcagag agcaggtggg 17040  
 agcggagacc ttagcgctgc gacgacgggt aacaaagagg cggagtgcag tggcgcgga 17100  
 ggagggcgcg accgctggag ggagacggcg agtcggagta gtcgagtaga catagtgggg 17160  
 ggcatatcgc cgtatggatt tatcctctcg gaggaataat taattaattg gtaatatatt 17220  
 atttttttta taatattaat ttattatata aaggtttcat cggactcggg attatgtaac 17280  
 ccgtcccgaa tatgattagg ttttataagg ttgattagg tttgattcaa taaaaaaac 17340  
 ttatccaaat taattgggtt agattggatc attggattat ttgtacataa gtttaagaac 17400  
 ataatcgatg gaatttgaca attcatgaag ggtatttagg caatgttttg ctaggttttc 17460  
 tttctttttt ttattttgta aaaaaaacat ttgtcttaat attttttcct taatgattag 17520  
 taatataaaa tataaaaaa tgttagcaac acactctttt taacacata ttttccattg 17580  
 gttaaaaatt attaaaaact acaaaaatta gaagaaaaa ttattaata agaaatgaga 17640  
 ctacacaaaa tttataaatt ttaataaatt ttaactaata atagtgttat taacatttct 17700  
 ggatatagag attaagtagg ggttttctaa tcttttcttt attaaaaaa tatgatgtaa 17760  
 aggtctaaat ttatagtctt catgtaaatc acgtgtgtgt atttgaaaa tattcaagta 17820  
 tctattagag ctatctatat aatacggtag agtggagttg agtcatatag cctctcacct 17880  
 acaatgacta ggactctcgc aaatatttgg gcgcaaacct ctgtctcttg tattttttaga 17940  
 gaacgagagg taatatatga tocatccatg agtatgcgaa ttgggaatga tgcgatgatt 18000  
 tgctcatatc aaagggggag aggggtgacg actaatatat ctatagttta aatagataaa 18060  
 atgttttgat ttttgtgata tttttttaac aataatcgtt gaatatcaa aacttattat 18120  
 tattttttta ggacaggaaa ggtgctgtaa atgaaaagga agtaataat tgaagtatat 18180  
 caaaaccaat gctttcggtt tctgaaatc gagatctaca taaagacaat tctaataaat 18240  
 gttatgaaga gtttattagg aaaatcaatg gcaaaaggaa gaaaatttat taactcgatg 18300  
 taattaaaaa tggtgaaata ataaaatcta tattaaaaa aaattacgca atagctcttg 18360  
 aaacgaatcc ccaacaacg atatgaattt catgtagaaa ttgtgtactt tgatacagaa 18420

agaggcaaaa tgaatgatt ttgtcatttt gagctttgat ttattagtag aaaacggttc 18480  
 caaatcttga attgccaat tttgggaaaa tggtaatcca gattgatctt ttgtgtgagt 18540  
 gtgggtgatct tgaatgttca ttaagagcta atcataata gtattttatt atcacacaca 18600  
 aaaaatcttt aaatatactc ctataaatat ttatatataa acaaggcact ttaaatgtac 18660  
 aacgtgcaca aatcatacag gattcgtagc agacctaatc ttctgtgtgc cctaggcatt 18720  
 taacctaggg aaccaaggat caagttacaa aagaatatta ttttaaaaga atttgttttg 18780  
 gactatatag accaagaac aaataactgt ggctatgtac actagcaaga taaataatac 18840  
 aaaatttaga caaaaatttc ttttcaacca tgtatcctgt tttccaattt caaagctgtc 18900  
 actaaagaca gaacgggta tcacttgta gaccactgaa attctatttc taaattcttc 18960  
 gaggtgccac ttttactttc cggaacaga gtgtactctc cagatactgc acctagcatt 19020  
 actaccgat caatctggat tgtcactttc ccaaatgaac tctgtgtgca gataaaaaa 19080  
 aaatagatgg gttaagcga agcctaaac caaaataact gaagatggat agactgcatt 19140  
 tacaagtcca atttttccac aattaccttt ccatttttgc tctgttttt gcaagaaatg 19200  
 tgaagctttt ggccttttgg aggactctca aaggacctg taaagctctc atccactcc 19260  
 ggattagggc cagttgagac cacctgatca gcacattctc atggtaatta gcacattcac 19320  
 ttttcgagta ggaatcaca tcaaataga aagaaaacat gattttccac gaattgtata 19380  
 ttgtgataac cacagattca gaatgcaaat cattgtccct tgcgtttaaa acaatagaaa 19440  
 agaaattgaa gtgggtgtgg ctcttggcaa gtgagaacca tatacttccg tatcaatatt 19500  
 gtaacaaatc gcatttaagt caaagctac ttcaaaagga aagaaaaaaa caatttgggt 19560  
 gaatctggat tctaaagtct ggcaacagaa ctgcaatgga atatgtaata gtcaactgga 19620  
 ttttttttcc taggaagtag ctataaaggt ctatataccta attgggcatt ttctaatt 19680  
 ggcataattt cgataaggac atagcagata acattattta ttatgcaaaa gattgacatg 19740  
 gagatacagt catataattg tctcacataa aacaagcagc tcatccattt caactaatcc 19800  
 tagtgcacta gtggagaaga agatcacaaa gcaagataat aaatacaaaa ttatttttaa 19860  
 aataccttgg tttgccttgg aggagtgttt ccgagtgtaa gcttgcaaaa aacacttggg 19920  
 tttccaacag attgcttcat gttgttaccg cgcttgatga tcaccaccaa tgttctctgc 19980  
 aaactctgca ataaaaattc tgccttctcc tgaaagcgag gtgggccaga ctggattaag 20040  
 tactgcagca aaggtatagc atctgcagct gctattgact gagctcttga aacttctgct 20100  
 ggacatgctg accaagcttg cctaagcagg aaaagtgcatt tcaaggcagc ttcttgagta 20160

gcctcagagc ctgtctttaa ggacgtaacc aaatggggaa tgctaagtgt tgctgggtca 20220  
 gttgctctca gtcttgggaa gttgctaagt agagaattta aggccttaag atactcatca 20280  
 ttccacagatc cagtagccca taagtccttt tcaatagcag ctaaagaaga agacacaaaa 20340  
 aatggttaggc atatttcatc actgatgaga gaggttatca aggtgaagca acttattgtc 20400  
 agacagacag aatcaacatt gctcaccaat gaattcagat tagaataagc acctttatct 20460  
 acagacaaga ctttcaaatt caacctatac tttctctgat acattaatga atatttacag 20520  
 aagactaggc gcctatttta gctagaagag gttatataga acaactaaga gtcaagctgt 20580  
 tcaacttatg aagcaattaa gtgcactaca agttttacta ttcatttgtt ccttccccaa 20640  
 atttttggtt tcattaccaa taatcaacca aaaagaatca gctctgctat ttttccaaaa 20700  
 gatgtattaa aagtcacatg tagcattttt gttttttttt tttattgaaa atcaagttat 20760  
 gcttctgctt ataatgcaa agaaaaatgc tttactatgt ctatgactat gagaatcaat 20820  
 tatttgattc aagaacaaga aaacaacatc agaaattatc atgtgaaaag taacaataac 20880  
 accattctat atgatgaaaa gcgaagataa gatataccgt agtgggtcag agtaaaaaa 20940  
 acaagtattt aacactgcac gataggaaaa taactcacca gttattgtct tgactgttct 21000  
 actagaagca tactcttgaa tggatgatt tgaaaataga agtttaataa acattgcagc 21060  
 ctgaacagat gtttcaggat cactgaacc tatcagatcc agtataacct gaacccacc 21120  
 cgctctgca actgctcttt tatttgatcg actgtacatc acaaggtttt gcagagcaca 21180  
 tatggctaca actttcattt cttcagttgg ttggtcttca agcacattca ctaaagcacg 21240  
 acaagctgaa actgcattac ttgtccgagc aagacctca tttgaaata gatcaccaag 21300  
 agccaaagtt gctaatagac ttgcttgttg tgcttgggtt tgggggtcca agagataatg 21360  
 ggataatggt aagatggctg acttggtaac ttttgtttct ctgatcttta cattgttaag 21420  
 caaaacttct aagagtcttg cagcagtttc ctgcactga tgagatctta ggagctcaa 21480  
 tagagcctct atagaccac tttcagccat tgcttcagcg ctagtccgt catcacttct 21540  
 caacacaaga agagcattta atgcaccaac aactgtgctt tctaagccag accgaagcag 21600  
 cctaaccaaa acagcaactg gaacttccaa gtatgatcca gaactaaatt gtagaatact 21660  
 agccaaaaca gaggcagcag actccataa agcatgagga atggaaggat cagattgcaa 21720  
 tataactttt gaaatttcaa taacaccacc ttcttttgca atttctattg gccatatcaa 21780  
 tgcaatactg accagggcct ttacagctct ttgctgcaat atatgtatac cagaaccaag 21840  
 aactcgtata agagggccaa ttacctgttg cgtcacgga tcttctgaa ggtgttcttc 21900

taaagtaga tgcgagagaa gctcagcagc caactgttgc actgctgata ttggagaatc 21960  
aagcaagggt ataagaggtt caataacttg gtgacaagtc aatgtgtaat cagcagcaga 22020  
ctcggtatgt tctaagatat taactaaaac ctgcaatgca ctatgttgtc catctggccc 22080  
aaattcttct cttgtcagca acataaacag gggttcaact acttttagcag cggatggctc 22140  
tttagctatg ctagcattat tggccaatat acgcaacaat tctgcaaaaag ctgcacatag 22200  
ataatcagggt gcttcatgaa ggatgtcaag tatgctttca ataactccag ctttcaccat 22260  
ttccatttta caagctggcc tgtcttttcc taacttgacc agagctctgg aaatggcctc 22320  
atgaagtaca tgattctctac cataaagtag gccacaaga ggaataactg caccatgtgc 22380  
agcaactaat tcagccagtt gctcatcacc aacaagccta tccaatgac ggacaactga 22440  
atggtgagca ggactaaact cacttacaag gagagagacc agagggtcaa cacaatgtgc 22500  
agcagccatt gtgaccgaa tccttgatc tccaaaaaga acagaacata attcagcagc 22560  
atcccccttg aggtccattg agcaatctga tgagagaatc ctgcaaagaa catccactgc 22620  
attcatctca acatctgcaa cagcaagtcg tttgaagggt ttttcaacta gtaacctaac 22680  
caatgcagca attgcagcat gctgctctct ctctaaacca gtattaagaa tctccaccaa 22740  
gggttgaaca gcttgcgag cagtctctgc atttctaata tggtcagcag agaataaact 22800  
ttccaaagct tttgcagccc tataccttgc agctcttctc cctaaacgca agacagctac 22860  
aagttgagta acagctccaa atgctgactc atgtctccga atttctgcac tgctaaacag 22920  
aatctctaac aaatctgtag cagctctctc cggtgcatct tgcggacctc gtgaaagata 22980  
cttggaaagt gcctctagggt cccctgctc aaccattaaa atcttatttg atggacaatc 23040  
tatagaaaagc tgagtcaaaa gcccaagtgc taaaaatggg gctcctgggc gttctggaat 23100  
tggtttgagc agatcaacta gagcaggtat tgccttccga gaagtggcac caactcttat 23160  
gtcatcaact ctaaaacacc tctcaagagc aacttgatca ggataatgca ccaagaaaaa 23220  
ttcatctgac aattccagaa gatcctgtat atcagaatca gcacagccaa gcaaggagat 23280  
aagtcacact gctgcccagc aatttgccac agacagaagt gttcccctgc taccattaca 23340  
gactaggcta gctattgatt gtgcagcaaa atatctgttc gctgattctc ctgacttcaa 23400  
taagttggca agggctggta tagatttcat ggttgcatgc gctcgatga tatccctatc 23460  
ttgaaataat actgccaaca acaaaagcaca aatccacata ctgctatctt cttttataatc 23520  
aatctgcata tcatatcacc ataaaggagc aaagatttac acacagaaat aatagttata 23580  
tactaatctg ctatctaact actatgtcaa tgaggagggg tgtatttgta aacaactttg 23640



aagatatggt atgcgtctat gcaataaagt aaatatttac tatggaaga atcagatacc 23700

actgtctatt tcaaatggac tgaatacatg cactagaata cctgactata ttgtgagaaa 23760

caatctgcaa tctgtcagt gagaacctca attgctccag cctccattat tgcaatttta 23820

cttttttcat catggcaagc aagaacagac agtaaccata tagctaagtt ggcactggaa 23880

ataagagcgg tgcagtgct tgatttaca tcatgtggtt cttttgtatg cctgcaaatg 23940

ctgataactt cctgtctgtc atcacctga ttatccaatg tagcctgtga aaaaataagc 24000

atatacgacca gagactgaac aagattagca cacaagtttg ataaattgag gtctccacc 24060

agcttctgat gatttacttt agcagacaaa ataagaacag cagctcctcc aattttgacc 24120

tttaccattt tagaagtgga attgattatc cttttagcta tagaagatat acatccagag 24180

gcagtaacaa cagtgtctcc cagaacaaaa ggctgatcct tacataatcg agataatatt 24240

tcaatagctt taccctgcaa cactggcgtt gaatcagcaa tagacaaaac tattgggatt 24300

atgcttttag gaaattcagc taaaactgcc caagcagggtt tactgtgtgc actggtcaaa 24360

tctgacctgg aaagcatggc aagtgttcca agagcttctg atgtggaact atgttcatca 24420

atagcaaaat ctaagaaga aactaatgca agaacagtgc cagctcgatt cacacaatca 24480

gttacagaat aatcaacttg cctcttagaa tgtaagaggc gagcaatttg tgcagcagca 24540

tgtgttttcc cagaaattgt gccttcacgt aatactctgg tagcagccaa gatgacttct 24600

tctgcaacag ctttctcagc aatttcacta tccaaaataa gatttgccac agcacatggt 24660

gccagctctg ccacttctaa aactgaagag ttagctagtg cagctaggga gggtaatgca 24720

tctctagcaa tagcagccat atccttgctc tctttgatgg aaagaaatat agcagccaa 24780

cagcgcgagg actccattaa gatactttca gattcaacat taagcaactt cattgctgac 24840

caaaagagtt taacagcaat gctgctttca cgcacatctt tctctgttcc aaaaattccc 24900

gccagagcag atgcagactt tgcctgagtt tcttctctgg ttgagctcaa tagttttatc 24960

attgtacaaa ttgcactact agcagcacta ccttcacgta gaagatcagt aagagcaaca 25020

actgaaagca tactttctaa agcatctaaa acataaacct tagattcagg taatacacta 25080

gtcaataatg ctgtaagctg actgatagtt gctgtatcag atttatgaat taagtgatc 25140

aaagtctttg ctgcaatctc cttccatttt gggcttccat tcttcaacag ccacagcaat 25200

gcaggaacga cttcagcact ttcaacacaa gcacgtatat cctcactgtg atcgcatagg 25260

ttctaagga ttgttgctga gtcttctctt gcttttgag atcctgactc caaaatttga 25320

acaagaggag gtataccacc agcagcagta atggccattt tactttcact attttcatta 25380

gataaaaggc aaaggagggc aacagcacat tcttgcctgt gttctgatga aagcccacaga 25440  
 agagatatca acagctgaac tcttccacgg ccttgaagtg cagcccatag actgcattca 25500  
 gtattgcaca atgtcaagag agactttaga agctcctctt gcacttcatt ggcagccatc 25560  
 gttattaaac caacaagcaa acgcttttga tcagagtgtt taagtttaat tgacaatata 25620  
 ggattgctat acaactagc tagggcttca atgggtccgt cctgcacaa gaattggcaag 25680  
 tgaggtttaa attgttcaag taatgtctgc tcaactacca aaggatctga tgcccaggta 25740  
 gattctgcct tgtcatcata tatcataaga gctgaagcta aagccccc aa tggttcagca 25800  
 gcttgagtag gagaggagca tgattcaagg ctttgaccaa gactggagat gacatacgac 25860  
 aaccaccagc agatgtttgc taaagcacac atagcatttt cctgtaacgc ttgagcatic 25920  
 tcaccctgca tgaactcttt ggaaggagct atggtagcat taatcaaaag aggaatgccca 25980  
 ttggaattgg ctatctcttt ccttgcctct ttgactggg cagatagaga tttaagagca 26040  
 ccagctgctc cagccctgac aggggcatca ttaccagggc ctaaaagttt gaggagtgtt 26100  
 tttgtgtctc ctgctgtcaa caactttgaa caaacagatg catcctccat catcatataa 26160  
 gcaagtaaaa agcacacatt ggctaaagtg ctgactgcc cagttgtcag taacttaatt 26220  
 aatatgtcca ctctccagc ttgtattgtg gcattccaga atctctcagt gctgtggag 26280  
 agatttttca atgccccagt caataaatta tcaactacat ttccagtctt caggcccttt 26340  
 tgcagctgct cccatagcac tgaacaact ccttcagttg aaaaaatttt tgatccaaca 26400  
 tgatcttttg caccaccttg agagacagca aaaatagtct tgcagcagc tacttgacct 26460  
 tcagcagaac tggactttag tagaccaagc aaaggaggaa tgcacacctc aagcaagact 26520  
 ttcaccctca gctcattttc ttacaaaagt gagcccaaga cagttgctgc ctgtattttc 26580  
 acattcaatg accctgaccg gagaaggga acaagtaact gaactgcctg agaatgagat 26640  
 ccaacagcac tgaagcatt ttctctcata tctataagct ccaataactg ctccaagaa 26700  
 tactctttct cttgcataga agatgagctc tgacgcagct gctcaatgca ttgggcaaca 26760  
 cttgctaagt tcccatctgg atcctccatg ctgctgctgt tgcctacgtt cctgcaaaaa 26820  
 attaggaagc taagatacaa aagtttgac attagatttt atgctcaaac tcaaatctcc 26880  
 atagaccata tatcaattat caaacatca agttagatag ttctcattat ctctatgta 26940  
 tgagactatc tgagggggtg agaggggaac attttttttc tgccttagaa cttttgacat 27000  
 ggataacaca ataagaggga ttcagaagtc ttactctagc tgcaaacaga agtactaatg 27060  
 taacagcatg gatttttagct gagcattata agcaattgat ttggtgattt tttaagccat 27120

aaaatattct ttttgacat aattaaagg catatcttaa gctgtaacca acctcaatcc 27180  
catcttcaga actgaatgag gaggtagagg ctctgaatct ttagccttcc catcaccatt 27240  
tctttcctgc aatcaaaaga atcacagtga gcaagtttaa acggaatgtc tgtacatctg 27300  
taaatgatga aagtaacaaa agttaacata aacatagaaa caatatatta gattcgtgtac 27360  
aatatgata aagatcaacc ttgaagttct ataagaggag aaatggaaaa tgaagtgtga 27420  
aaggaaaata gccaaacctt atatacaacc ttgacttcat tttgtatcaa gtaattgaag 27480  
attttcaaaag tattaacac acatgattgt tgcttccggt acagttgact aattatcatt 27540  
aaggatagca tcttattgca caaaacagag gggagatata aaatgaata aaactatac 27600  
agcattaagt aatgcaacaa agcatttgta cattacacta gtgtagaaca attaccataa 27660  
ctaaccatt aagatagatg gtcatgaatg ctatattttt aagagaatct aataatgcc 27720  
ccttttattt ttcattttct agcattacc tgtactcaat cctagtcca aggatatgta 27780  
taagcatata ccaaatgatg gtagagtga tgactaaaat aatgcttgct acctaaaata 27840  
gaatatttct catcaacttc ataatttta taaaaccatg agcatctcaa ttcaagaat 27900  
ggcaacagat atgttccaat gggatataag aaacagataa cgaatgaaa tcacagagcc 27960  
tctttagatt tccatgggtc tcatcggtac tttctagaat tatttgatac aaaaaaaaa 28020  
aaaaagtggt taatacaaat caaggggtgt tgggctagtt ttagaagtct ccacaaccac 28080  
aactgccgtt gcactccct tccactgtct gcaatttcta aaatcacagc ataacctga 28140  
ccgaatttta aaacataaa tgaccagcaa gcaatctgaa gcgataatat atccaatgtg 28200  
tacaggccga gttaaggatt cataacagtc aaactaatt gtttatacta attctcaaac 28260  
acttttctc caccagtaat gatgatgtga tgacataaaa tgaataata aaaaaggaga 28320  
aataatggaa aaacatgaag aagtaccata tcattagcag caaggggtgt gccattattg 28380  
gcagcaaacc tccaagtcat tgtagtgcgc agcttcgctc cgcgtcact gaattctct 28440  
ccaagcagat caatccagga cactgctccc tctccagatc tacactggag agttcgtcag 28500  
cacatcaaag cctccctgtc cacaccacac ggtaattcag atcagaaacc aaacagcaat 28560  
tgcacagaaa tcgcaaatgc aagtcaaca agttaaaaa tacacagtga cctaaccaaa 28620  
acactgtaaa aaaaaaatc aagctaaca tggaaatcggt taaaattgaa tcgcgcgaga 28680  
gacagtca gaagattggag gtagcagatc aacagctact accgaggaat gggagtggaa 28740  
agcgtgcagc ggggtcaacg aaaaagtcaa agcggttacg acggaatgca ggaatcgaat 28800  
gaatgctagt gaaattgaa gaattgaaa aaaaaacgaa gaggtgagag agaccttgag 28860

ggagtgagaa gcgaagcag tttgttggtg tgagaatgac gcatggattg gtttagcgg 28920  
ggcggtggtg gtagtgggag acagatttga ttattgatta ttagaggatga ggcttcgtta 28980  
agaaaagagt gcgttatgtg atttttcggg tttgaattga cggagaaggg agtgtttgtg 29040  
acaatgagca ccgacaacag ctccagcctg caagctcata gcttcgcgcg gtcatttaac 29100  
ggctaggtc agccttcagg ttcgcttgag ccaccggtt ggaatagaat cagctggaaa 29160  
taaaataatg tacctctttt tacattttta tatctttct cgcattggaca aatttgtatc 29220  
tcctctttt tcttttttaa atttttttct catcagtaca cttatcattt ttattattat 29280  
tttctcttct aattccaccc caatctattt catttcagag tagataataa tcaatactta 29340  
taatcttttg tatgcatatt tttattagtt ttaatacgtg tgcagacatg tacaaaaatg 29400  
attaataagt attataaata ttatattaat atcgctgatt aatatttact aacataaata 29460  
ttataagtat tataggatg cgttatattc atattcatat ttgtatttta ttatatatat 29520  
atatatatat atatatatat atatatgtat acatattata gtaaacccac ctcatactaa 29580  
tgataggaca caaatcattt tttgtgtttt atgtattttt ttaattattaa ttaatttcta 29640  
ttaacttta atgtaaaaaa taaaaaaaaa tgaacttttt acaattttta agatttttga 29700  
taccatgtga aaatcactat taatttcatat gtttcataaa ttgtaagttt gttgaaaatc 29760  
actctcttaa aattttcagt gattgaattt tctctttgaa gatttttttt cttctcttaa 29820  
attataaatt agatgggtga ggttattaat tttggaatc acgtaattgt cctttggaa 29880  
taccgaatgt cttttttgta ttcatgtttt tcaatctcaa atacaataa atagttgott 29940  
atatttttag ttttttaaat tttaaatttg cgtaaagaaa aatagattct ttgaaatgat 30000  
ggttctttga aagtatatat attttttata tggcttgcat aacattgcga attttaactt 30060  
gtgaagaat taatgtat acataaataat gtttgcctaa gaattttttt tattaataac 30120  
gttaaccaac tacttcaatt aaagttttta tgttccaact aagttcgtat atattaacga 30180  
taogaagaac tcgagaattt tattgacctt tttttttttt gctgagatta ttgacctat 30240  
gtttatgaaa taaaaaatga taccagtgc aacattattt tcaccaattt cttacagaaa 30300  
aacgtggtat atataactg gccattggaa aatagatgac taactgttta tacagagttt 30360  
gttagtatta agttttgtca aaataaagtt aatgcaataa ttatttcttt atttacagta 30420  
gtttttgaca tatacaatat tattatatat ttttttagga aagaatatta ttatacttaa 30480  
gtactaacta ttattttctg gcaaatatat taactagtgt tatttatttg tgagtacact 30540  
cgattttaat agaaacatta attattaatt accttttctt tgtacagagt taattaacat 30600

tggatatata cttatatgag attttaattt gaatttaata ttttttatat tcagatatatt 30660  
 tattttaatt taaaaattga gattagacat tcacatttac aaaagataat cattttcata 30720  
 atagtttttt taatatatat atatatatat atatatatta ttaagaaaat 30780  
 ttaaattctta tgcacaaatt acatatcagc cctgctcttc ccactaataa aaagcatata 30840  
 tgcattcggt attggaagta attaaagatt tttgtgattt attttttcca aatttgagaa 30900  
 ttttagattt cttttaataa ttttttttag gttaaattaa tttttccctt tattttattta 30960  
 gttagatcaa ttcaatcctt ttaatttttt aaattgggtg aatattgtcg tttttatccg 31020  
 atttagttaa aatcaatagc tttgtaattt agtcccttag ttttgaaagc attgtaattt 31080  
 acatgacaac aataatttag taaataaatt atttatccaa ctgaaataaa attatcttaa 31140  
 ataatatata ttagaaaaaa atgttgattt actattataa tttttttcac ttgtatttac 31200  
 attttttttt actagtaaac tgtttactta aaagtattt tatcaccatt tacagataat 31260  
 aatatttttt actgaaaaaa ctttacttta atctcatttt taattaaaga caaatttata 31320  
 aaattaaaa gatttgaaat catatataca aaatattgat ctgatccaaa ctcaattgat 31380  
 tttttcataa aaaaacacat tatatatgtt tttagcccaa taattaacaa gaaccaaact 31440  
 ccaataatc gacttaaaaa aaaaacttca attattcggt gcatatgacc acatgggttt 31500  
 cattttagct gagcttgagc tcttgattgt ggccaagcca tagcctacat ccaaaaaaat 31560  
 ggaaagtcat ctcaattata tccatttttt ttgtgcttag ttggcaaaag tatattttct 31620  
 ctctcactct attgcacagt ggagcgtttg atggcaatgg aatcctttga tttttcaaag 31680  
 gtagtaataa aggcataata gaaaatagtc ttttattata agctgtgaga ggtgggtgca 31740  
 gaagttaaag acaagctacc aagcattcat gttattacta gcatatatta tatgcaccag 31800  
 ctggcaagtc atgaggctag attattatgc aataaaatca atgtgggtca tgttattcgc 31860  
 acttgacccc tgtaagcatt tgcgtttatc atgcaatcat agggaccctg tacagtacca 31920  
 gctagctaaa ccccccatta cttgttttat tattctagtc ttaaattaaa ctaacagcaa 31980  
 ttaggttgca tattaatgtt aatttccttc gatcgatttg tctaataccc tatactgtct 32040  
 tgcccttgtt tcatttgcac acatccttat tttttttttt tttcttaagc agatccagat 32100  
 gaccaagtgg cttttgttat ttcattccat ttctcaaaat actgattcaa agtcgagttt 32160  
 ggctctcttc gtttttttag ctctgtttgt ttgcgaccac attggtaatt tgggtgcata 32220  
 gaaattgagt tgggcttgtt tgggtataact aatgcattgt agaaataatt tacgaccsac 32280  
 ctaagcggtc aataattatc attataacac taaattattg ataaacattg tatttgggctt 32340

tcatttaaga gtgaatatgt gatgaattgg gtcattcatt gtgtagctcg ttgaaaaaat 32400

ccttacttgc tactagttag ttaaaagact aacatcattt ttattttcag ttctatcccc 32460

atgtaaaagc tctctagaaa atgggcaatt tactctcaat acaataattg agttctctca 32520

tgagttgaag agccatgaac ctaataataa atctaggtat tcaggtagac gtttattatc 32580

ccctctttgt cgtaattaat atattgtaat ttgtaccctt gtgtgatttg gacattctac 32640

acacttgctt taaaattcat caattggat caaaacaact tagatacaaa ttatggctat 32700

tgtttttcat tgttctttca tcaaaacatt gtcaaaacgt gggttttttt ggcttatttg 32760

ttatttttct aagattatta aaaacttgt taatttttca tacattaaat ctataattgt 32820

ttgagggttt ttggtttatt tgcttatttg ttgtgttccc ttatcatgtg cgcgacaaca 32880

tattaaaagt gtgatttagt aaactaatta taaagaattt gatggttggg ttgcaaata 32940

tccttccacg aaccctaaac atttcaaacc aaattactca cagatatagt tttttcaca 33000

aattgattga gagatcgtgt agaaaaagta aatttcaagc aaaactcatc ctcaatatcc 33060

tttttgtatt ggtttataaa catgagagaa ttgggaaaaa gcttttgttt ttgatcoat 33120

tctcacatct gcgattttgc ttctttttct ttctttcttc catgtatgtg ttccaaacct 33180

ttcttctgaa attgattggg gattctctat attgatagtt ttcaaacaaa atttctcaag 33240

ttttactatg tatagggaca ttttatcaaa tcataaccc ataaaaattt attttgttgt 33300

catttttgtg aaaaatatgt actaaatagt ttatctgatt acatgtctta gtgagtgtct 33360

gatttcgatt aaaaaacata tgtatttgat atttttcaca acaatttcag actaaagaca 33420

caaaagccta atatacttag tttagtgtag tgccatcac cttagggtgt catacatcca 33480

aacacttaca cacttgtttt agtaataat gttatattct atctcaataa aatccaatgt 33540

ttagaataaa tccttgattg acttattatt attaatatgc tcatttataa attattgtca 33600

aagataattt tgttgaaaca attttttaaa tatttcaaaa atatacgttt ctgcgttgtt 33660

atatgaatca tattgattta tagttggagt ttttagttga tatcacatta catactttta 33720

gtgtgtgatt ggatagctcc aaaatcatgg ttgatcatga tgaggtaagt caatcttaat 33780

ttttatcaaa tttcatggtt tagattattt ttaaaattat ggtagaatca attatgatgg 33840

aaaaaaaa catggtagaa ttgaacaaa taggagtaat ttcaacttct atcaaaaatca 33900

gttatgtaaa atcatcatca atctaacac acacttatca tgttatcaaa gctaactgat 33960

acaaaaaaat tataaggtag ttttaaatc tatggcccg actaatattt ataggatttt 34020

gctaactaat atccttaaaa cattagttaa aaaactaaaa gaaaaaaat acttatgatg 34080

gagattataa gaaagcataa aaaagtcata aatgataatt ttgtacatct caataatttt 34140  
 taattcttta atcttttaaa taatattcta agaacactag ttaacatttg atattttttat 34200  
 agatatatta agtggtatta atattttaca tgtgggtatgc ttaggcctag taagcccttc 34260  
 aagatattat gcataattaag gactttcttt tcaaaagaaa agtaataatt atatctataa 34320  
 aaatatcaaa tgttaactcc gaaaatatca aatgcttagg cctagtaagc ccttcaagat 34380  
 attattcaat ccgattgttg tcctctctgt atggtaacta ggggtgggaa taggccagac 34440  
 cgacctacat ggacctcga cctgcctcac ataaagtctg gtctgaactg acccatttaa 34500  
 ttaagagggt aggttcaagc ttttttaaaa agcttattaa attaaataga tcagacttag 34560  
 gcttggttaa aagccttata agtctgatag gtcggcctat atatatatat atatatatgt 34620  
 atatgtatat atatatatat atgtatatat atttatatta ttgtttgata ccaatttata 34680  
 cttatattat tttttgggta caattaattt ttttttgaaa ccagcagact ttgattatac 34740  
 attactgtct cataacttcc attcctataa tcaagtaaga ctttaattac catttaggtg 34800  
 tgaattatgt acccctttat attaccctag atttttattg gcttttttgt taacgtttcc 34860  
 ttttccatta acttctcat tacgttctca cgtcaaagca aacaaatata aaaaggctat 34920  
 caggccaggc taggttttta aaaaggtcag gccagagcaa aataaaagct ttgtataggc 34980  
 tataggccag attcaggcct caaaaaattaa tcgtaggcta ggctcagggt ttttaaaggt 35040  
 tggctatagc tggcctatcc ccaccctaa tggcaactca ttgagagcgt aattgggcca 35100  
 agccacggcg attgttattg tgagtggcaa tttgggggca attgcttctt acacctctaa 35160  
 attacttaat gcaatcccat aatttttaaat atatccaaat tgacctattt ataacttcat 35220  
 ttctaccttc tttctcattt ctctctcgtt gacgcattat tgagggtgtc agaaggagtg 35280  
 gttcgtatac atttttagga aaggaggtgt gaagtgggta ttttaactat ttttttccaa 35340  
 ttgtcatata ttttccaaat agctgaattc ctaagataca aaaagaaaaa aaaccacttc 35400  
 caaaaaaaca tttccaaat catgttaaca ttattcctaa attgaatttt tggaatgoat 35460  
 tttttctgga aaaggtaacc atgaattatg ttagcataat ttcataaatg cttttttgga 35520  
 agtgtttttt tttgttttt ttgtatctaa caagctcaac tatctgaaag gtattagaga 35580  
 tttttttgga ttacatattt tggagaatgc ttaaaatcaa caaccatacc agtcaattca 35640  
 ccaacaacat gccattgtcta tcttctgtcag gtgcaaaaag aaataaaga agaagggggt 35700  
 tgaggaagggt ctgcgctccc aaagggaaaag gaatgaggta taagagacac tcaagggcat 35760  
 tttaaccat tcaaatgttt ttaaaattcc tatgggatgc aggaagcaaa ttaaggggtc 35820

taggaagtaa ttgtccaatt tggttttcca acaaaatcat gtttatgggt aatccactca 35880  
 ataaatagat gcactagcta aattttgttt gtctcaatgt gttccttttc atctctttca 35940  
 ttttgtttct catttttcat ccttaccata tgtttgctaa agtgctcttt tttctagaaa 36000  
 ctttttagttt gggctaaggc cctgtcatga ataaaaatag aatgctcttg ggggtttgct 36060  
 agtgggtccca acctctctat ggaaaagtat attttgcccc tcttctcttc tacttctctt 36120  
 tctctacca tctcatccac tccccccca cctattatca cctaaccgc accagatcta 36180  
 gaatccctat cccatttctt ccatgccga caccctacc agccttacca tcaactattgc 36240  
 acaaaagtcc ctactcgtgt atgatctcca cttttgaggc tccatgggcg caactcctcc 36300  
 ccatcataac tatgccgcca tgaacattca cagtttgca ctcctttgtc ataattggaat 36360  
 cacatttttg catggatatt tggctgcata aaatcgtgat ttctgttgat taagtggggg 36420  
 tgttgcaaaa tataacaatg aaatcatgat ttgtttgtat atttagcgtg gaaatttttt 36480  
 tacgataaaa ttgtgattct attgtgttag gggggttgaa aactaataaa gaaaagaatc 36540  
 acaatttcat tacaattttt attcaataaa atcacaattt catataaagg atattcttgg 36600  
 aataaaaaaa cttcagcccc ttgataggag tcaagaagca attttaatag ggcaatagta 36660  
 acttccctcc tgtgttgggt attagaactt ttatatgtcg gcccaataaa gaaacagaag 36720  
 cccagctcag gcacggggag atgaaattct cgattatctt tcagccacgt gttaggggca 36780  
 ttatcggaat cacaacgaga aaatcttaat gtttagcaacc gctgtggaac cgagaccac 36840  
 ttggtattac ttcagggttc cctcacggaa cggaacaaaa ggcgcacaca atcgccaaag 36900  
 agaaaaagaa cctcaacca aacaacaaca acaacaaaaa aaaagcccta gaaaagattt 36960  
 tggagaagca atttctcact cctgaaaatc tcatctcttt cttctttaca ttgtttttgg 37020  
 ataaaacatt ttgctgatac tgaattttat tttctcttgt ttctgtttct ttctgacctt 37080  
 tttttttttt tgcttcgttg ttttcaggga gaaaggata ggagtaataa aagagagtgc 37140  
 attcggtttg gttcttcgac tggggactcg aatcttcgat gcctctctgt tttttctctt 37200  
 ctggcgcttc tttttctatt ctcaaattag ggtttcgtg agcgttctct tttctcgatt 37260  
 ctccgcaaac tattatcatt gtgagtaatt ctccgttctc tttctttcgc ttagtaattc 37320  
 atatatgaat gtatttttat atataagtta tttttgtttt gaaatttcat ggtttcctta 37380  
 gatcttgcg ttatcaattt gatgtccagt tctctttctg caccaaatc tttcagtgtg 37440  
 gtttaaatat taataaatag tttgtatgtg ttagtatatt tttttcatcc gacccagtt 37500  
 cagtgtacag tgatttttag atgaacactg tgattgttgg aattgattgt tgtttgtttc 37560



ttggtttttg tgaatttgat tgttggtctc aggttaggga aggtgatttg cgtttggaca 37620  
taatgtcgtc aaaatatcag attcttgaca accgcctat taatcagtgag aaggttacag 37680  
agctaaaga tgagctgaag aggaggaagt taagtatcaa ggggttgaag gatgatttgg 37740  
tcaagcgttt ggatgaagtc ctctgtcttg aaagggagagc tgacgaggct tctgagaagg 37800  
atgaggccaa tggtttcgat ggtcatgtag atggggaaaa agattcagag gcggttactg 37860  
tggatgctga aatggttgat tcaactgaca gggacaatgc caaacattt gagactgctg 37920  
aaaagggcaa gagcggtgta gttgatcctg ttgaaataga aaatgtggaa aagattccgg 37980  
aggttggtga tcacgacagc aataagaatg acaagcagga tgggtgttacc aaccaagttg 38040  
acattaataa tagtgtgtca gccatagatc aagaggttga acctaagggc ttgccagctg 38100  
gtggagatct tgcaaatgtg ggacaggagg caattgccca tggttctact gtggagaaga 38160  
ctaccattac agttaatgag agtgtagtaa cagaagtagt ggccagtgtc gaggattctt 38220  
atcgtgcaga aaagaataat gaggattcag cagcaaagct ggagaatgag gagtcaaagg 38280  
cgacagctgga tgggtgaggac tcaaaagccc agctggattg tgacatcaag ccttgcagtg 38340  
aggatctctg gccgactct tctgttcagg aaaaccaggt atctgagggt acccctagtt 38400  
tagggtctca agtaaaatct gattctattt ctactgattc tctgtcaatt aatcaaaaga 38460  
atgaactaaa ggatactata attactgata atgtcaaat agaacaagat attgttaagc 38520  
cagagatggt ggaagaacca tcatccagaa atgatgtacc tgtttcttat gatgaatcac 38580  
attcaatgga tgttgagggg ctgcatgaga aaaaggcatc tgttgaagaa aatttcaata 38640  
atgttttaag tccagacatg aacaagacca atagcagcga tgatgtgggg tatcccgaag 38700  
agttgaactt ggacagaagt tctggtgatg attctatgga agaggatttg cctgagacta 38760  
agcaaatgga ttctaagttt aatgttgatg aacttaaga caaaatagag attgaagagc 38820  
ctattgtgaa ggaggaaagt agtactatag ctgttggaag tggctgtctc gcaggagaaa 38880  
gtgatatacca ccaagatatt gacattagtc cagttgctcc aacagagaaa cgaaaattta 38940  
atggtaagca atatatgtgc tgaattgac actttttttt ttattggttg tactcttgaa 39000  
gggcaagaaa gaattatagg cactatatgt tgcttactg gcatgctgtt attatttctt 39060  
ctttgtattt ttaccatcc agtattgttt ggaacattt gtacagaaac aagcatcagt 39120  
tgggaacaat gagcctgcaa aaaggcaacg caggtggaac actgaaacag ttaaaggctc 39180  
agatgcacaa agcactactc ccaggcctgc tactacacct agagatgagc caattgcttt 39240  
gaaacgcaac ttctctaggt ctgattcctc tgcaactgat gatacaccta aagaacgcat 39300

tggtaagtca acattgttgt gattagggtta aatctattaa gaactaatgc tcctgtttgt 39360  
 tgtgtaattg tgggtgtaat ggagatcaaa tttcaattcc tagctttgtg gctcaataaa 39420  
 ataaattcct ggcaccttca gattttaatc ctggacactt cttgaagcac aagttgcata 39480  
 tatcttatcc tctaaaatgt atttgtggag gtttttattt taagatggac aactacctgc 39540  
 tgggtgattt tttcatgcac ttgtgcattt actgagataa gacttagttg ttgacttgggt 39600  
 ggttgtcatg cgctttttta catgggtgta tttgtctat gttgtacagc atgttttttt 39660  
 atggattact ttgttgtcat agtggcatgt gttgtattgt ctgtctctgt tttcaagtt 39720  
 aaatagtcct tttttaatct ttggcagttc caccaccaca aaggtcccca actaattccc 39780  
 tcaggattga tcgttttctc cgtccattta cctaaaagc agtgcaagaa tctcttggtta 39840  
 agactgggaa tgcagcagc ttctggatgg accaaaataa gaccattgc tatgtgactg 39900  
 taagataact cttattgctg tcttcttca actgtaaatt ttcattgtgat gttgttatat 39960  
 tatgacacgg tcttgacatc taagtaatct tttgttgatt tcttggtgaa tcagtactcg 40020  
 tctgtgggat aagccattga gacacggaat gctgtatata atttgcattg gccaccaa 40080  
 ggtgggcgtc tcttagttgc tgagtatgt gatcctgaag aagtgaataa gaagttagaa 40140  
 cctcctccta ctacggtgc atctgtcagc actgtccagc cagttcctcc tgcacctccc 40200  
 tcacagccag agccttcccc tcgtctgcac agggagccgc atccagtccc agctactctc 40260  
 ccacctccac caccattgtc aaaaccccca ccagtagcaa gagaacgggt tccatctcca 40320  
 ccaccccttc ctgagaaagt tgaccacccc attgtcactc tggatgatct ctttcgcaaa 40380  
 accacagcca ctctcggat ctactatcta cctttgtctg aagagcaagt tgcctcaaaa 40440  
 cttacagcac agggtaaaa ggcgaggcag tagataggcc attagcctgc aaatcttctt 40500  
 cccttcgagc ctttactgat ttggatattt taattagtgg tatgggatta ttttgaatgt 40560  
 cttggcagat ctgggtgtgt tttgtctcgg gaagaaattg cattgcttga gacttaaat 40620  
 tatttagttt tagaagtcgt gaactaagag ctaccggggt tggatagatt tteccactg 40680  
 ttgaaactag gcttgagcga acctttacct aaaattttat ggaaaaattt gacgttagaa 40740  
 tatgtttata ttcctagcca cgaattttgt tttcggtagc tggctttttc ttgatccgtg 40800  
 tattttctta tgctgagtgt cgattgattt tagtctgttg attgagcata gtatatttct 40860  
 tctagccacg tggtttgggt gacatcaagt ctagcaccgt aggaagggtg atttgtatgg 40920  
 gttgactcgc caatttatcc accgtttggg agaattgtat gcaagtttgt tgttattttg 40980  
 ggtacagtat ggcattggat tttattgtga tttggttttg atatgactta tataagtttt 41040

tcatccataa aaaatattggt gttttcaaat tactacttaa tcacgtcact gaaatgtgac 41100  
 taatgggtata acttgaaaca aaaaattttg aggttaattat ttgataaaaa attttttttt 41160  
 agataataat ttaaaaaataa tttatttttg aggtatgaaa aatttattta aaccttttga 41220  
 ttttggttgt tattgtgaaa atctaactaa attgcatatt aaatttatat taattttctt 41280  
 ttacattctt atttaaatta tcttttctat aacttaagct ggaactcatt ttttataatg 41340  
 atagtatgtg gttttttttg tattgaaatc ttagttatat tgcattattt atatataata 41400  
 ttgatactgc ataattagtt aactaatgat ttgtaagtta tttttttatg aagtttataa 41460  
 taatttaatt gttgtcaaaa accacttaag ctgaatgaaa actaaacggt gaaaactaaa 41520  
 aaaaaaaga acaacacggt gaacatccct actactaagg acgttgtatt tgtttgttga 41580  
 gagttgtggg cctgtggctc cctagtgaaa tcagtaaaaga aagactaact attcatttcc 41640  
 ctccctaaaa gtgtttaaag tccttcaaag atagaaaatt aatttttagt ctctcaatgt 41700  
 ttaaaaagtg taataaaatc attctataac tttttctgtt aactatcacg attttacctt 41760  
 acacttttcg attaaattgt gataaagaat gtacatatgt gttgtttggt tagcaaaatg 41820  
 agtaagagac ttatttgtca gccaaaagtg gatggattat tttgtctctc tcggagattt 41880  
 cttctatttt tctcccccc atttctttat cgttacaatc ccagcttccc ttccaatttg 41940  
 ttttcacgca gttcatcttt ttcacctgga tgctagcaaa tagatctggt tgaccctgta 42000  
 actgaaccgc acgctgcaac cattttttggc gaccaaatgg agtgggaagg gatgcatttt 42060  
 aatcatcttg tgaagggggg aaaaccata accaattttg caaagggttt gataaaaaaa 42120  
 ggaaaaaaa aaacacaaag aacaacatc atttgacaaa caatacgtg gatcactttt 42180  
 ttttggtcaa ttctttcatt tttgaacaaa taatcaagtt ctaatatatg acttgctaag 42240  
 aattggttta tctacaaaag cccttgattg atttaaaaa atcattgagc aatgagcgag 42300  
 aaatttgaag aaatgaaag aggaacacgt ttggtttgag caacaatttc ataatgagt 42360  
 gttttttgca attactcaag ttttaagtat gggtttgctt gaaattgatt gaactacaag 42420  
 aacactcaat caattttaca aaaaagaatt caatggagca aaatcaaggc aaaaaggga 42480  
 aaaaaaaatc ccttttaacc ttatatgtta ggagtatgct tttgggatcc aatttttaac 42540  
 ctgttttaac cttaattgca ataagcccat tggtagaga agaagatata gggaagagga 42600  
 gaaagagat gatggggagg aagttgggat agcaacagta aaaaatggg aaaagaaaaa 42660  
 aatacagaag atcttcaagg gtaacaaaat agttcaactt ttgggtgata aaaagttcat 42720  
 cctactcatt cccccgaeta agcaacacac atgtacattc ctatttggtc aaaagatgct 42780

ggtatatacaa aattgtgaga gttaaaagta aaatcatttt gtcacacttt tccaatatta 42840  
 aggactataa aaataatttt tcacttttca ataattgaaa tgcagaatt tgacactttc 42900  
 aaagactgaa atgagtattt atgataaaaa aaaactaatg tcacccattt agaatacaaa 42960  
 tttagcatg aacatgaata ttcttaataa aatattgttg caaagtgtgt agattgaatt 43020  
 tgagaacatc acagggttac accacaagaa aggatcaggg aatttttagg aaaaatttct 43080  
 tgagtaacca aaaaatgaaa aaccaatagg tgtagggata gtttgggttc caacaataga 43140  
 gtgtcaaaac ctgatagtgt ctcaagtcaa tgaatcttta aagcatctcc aataactttt 43200  
 cttacaatct aatctttcat gcaagatctg atctgtttaa gatttattac tagagtaaaa 43260  
 ggattgattt ccaccgtgga gatctgttct tctacaagaa taagcaagat ctgcatgtgt 43320  
 tttcaaaaaa tgatattata atgtctctct tcataaaaag tgtaataaaa tattagaagg 43380  
 tgaattcatt gtaggaccta caaaaagtg tttaaaatca caaagcttga ttttctattg 43440  
 aagcaaaaaa ttggaaagat ctctagaatg atatagtatt gcaaggtgca ccaaaaagtt 43500  
 gacttagatc acaagtttg atcttctatt aaaaatgtct ttagcgtgtg tttgatttat 43560  
 aatttcaaaa gttaagttta aggtaacata actttttaac ttaagtttat agataaaagt 43620  
 aagttgaag taacttgttt tatgtttgga tactttgttg taaaagtact ttgacagaa 43680  
 tgaaacttat ttggattatt tggatgaaa gtacatttga tattgttgca aaggacaaa 43740  
 tggatgttga ggggttataa taaaggtcag atatggcttc tttaaatttt ggcaggtaag 43800  
 gaataatgaa aataattatg ggtttaatgc tggttgttta gaggtaattg ctggaaggag 43860  
 ttattgaaat ttttaaaaaa aaccttatct ttctaacgga tccaacgaac gttaacggaa 43920  
 aagtaaccta tagttgcttc ccattaaaaa tcggatcatg tgcaattcaa gtttagttga 43980  
 accctcgttt gtaatggtct cacagtaagt tcagtgcac ttctactccc gtttagagac 44040  
 tcaaacgaaa aaacaaacac aactttattt tttttttcat tactgtcggt tgtacaacta 44100  
 ctatgaaca aacatgaaa atgcacaatt tactcgtaga aataataata ataataaaaa 44160  
 taaataaata aataaaataa agtatttttt gaaaaataaa agattgttag tactagttag 44220  
 tacggagtat cgtattcgtat tcatccacc tcggcggttg aaacaatttt catggttttg 44280  
 cgccaaataa aataggcgaa gaaccctaac ttaaaaaacc tcttctcttc aagtaacgag 44340  
 tcttctcttc ttctcaact aattgcattt aaaagtatcc tcaagtaact aagaagcaa 44400  
 tcaagaaaga aagcttcttg tattcattca ttcatcatt atgtgtgtgt gtgcgagtac 44460  
 aaaatttatt tctggaatct gaggtgttcc atatagccat gttgcttact tgcttttagta 44520

tggttaattcg gttaacaatt ttttgggtaa gtagcatagc ttgtttttac tattctttta 44580  
gcatattagt tcttcgggtc aatggaagac ccattacacc cttcaacaga aaattgcacac 44640  
accgagatgg tggaacagca agtaaatgta aacttatgtt ataagaataa gccactgaag 44700  
ctagataatc agccaccaca ttagcttctc tgaaagcacg aatttattaa ttgttaacag 44760  
aatacaatga cattatttga tccatgtagc tggctctacc tagtggaata ggcttctgtt 44820  
tggttttgggt gagtgtatta gtcattaatg caacgtaagt tgttgtcttt aacgcttgat 44880  
atagctttat cacattgggt tataattcac cggagcaaag aagaactggt tgttgccact 44940  
tggaataaac aatttgacaa atcagagatg attcaagggt tctctctctt gtatcttgca 45000  
aatgacattc tgcagaactg taagcgtaaa gggaatgaat ttgtgcagca gttttggaag 45060  
gttcttctcg cagcacttaa agatgttatc aagaagggtg atgacatgg aaagcgtgtg 45120  
gtatctacat tggtaagccc tttttactac atgtgatgca agtatttcaa ctgaataact 45180  
aaatgggaaa ggggttaaaa gaattgaata taaaatcata gcagatagtt cttcatgcat 45240  
ggctgaaact ctgatattgg tgtataaatg ctgatcatat gatttgttta ccatgatatt 45300  
gatctgttaa cttagttctt tactaacttt agattgaaat atgggagcaa aggagagtgt 45360  
ttggatccca ggctaggaac cttaaagatt tgatgcttgg agaagatgca cctccaccat 45420  
tggaatttgg caaaaagcga tcacgttctg taagaattgc gaaaagggtt tctcactcca 45480  
tcaaatcggt gagactgagt tactgagggc ctttgcaact acatccttcc tctccccatt 45540  
tcatacttaa atcattgatg ttagatagtt acatccttca tagaggtttg ctatggttct 45600  
ttcatatctc acttctctga tattgtcttg atgccagaaa ctgtccatga gaggtacagc 45660  
agaaaaata gtgtctgcat ttcatattggt gtcagttag caatccgctg aagatgcaga 45720  
gatgagtaag tgcaaatctt ctgttcacgg tgtgaggaag ttggaataag atgttgatag 45780  
cgctgtctct gttggaagc atacagcaga acagagttac agaagtaaca ccttaattat 45840  
ggaagttaga gttctaata tgtttctttt ttccagacaaa gatcctaaga gaaatacttt 45900  
agcaaaaggaa ctgagaggag aggaaaatgt tttgaaacaa tgcatagaaa acttaaatata 45960  
gttgagcaaa gcagagttag ccttgatctc cagttaaaag aagctttgca tgagcaggta 46020  
agttatacac ttgtattcaa tattttcagt attttattga gttttcggtt gttcctaact 46080  
gtgtcttcca attgtttgtg gatttttgta ttaggaaatc gaactagaga atgttcgtac 46140  
acagatgcag gtacgtactt ttcattgattc ttttccacag aaggttgta tttttgttc 46200  
atttatcttg gtttgttgcc atttaattac ttctctaatt gtacctcttc tgcataatc 46260

cctaggaagt aattctgaat attctccttg tagtttattt tcatgcatat gaagagttaa 46320  
ttttcatgaa ggtgttcatg taggtttaat tatttctctt atggtacctc ttctgcatat 46380  
aaagagtgtt gtctggaaga tgtttaata ggaggagatg ctttgcctga tgaaatttaa 46440  
atggggaaga tcagtgtaat ttctgaaaca agataggggtg ttgcgttaa gtttcataaa 46500  
ccttgacgga ggttagtgta ataacttga tttatgaaga ataacagttt ttggttaaag 46560  
aattgcttag tgaaagcacc gtacacaatt tgagagtaaa aaatctcttg ttaatatattt 46620  
aatgcccggt attacttggt gagtctgttt tcattattgg catttggaca tgtttgcggc 46680  
tttttagcta atcgcaacga caaatttgaa ctttattatg cttaaatggt tgtttaagga 46740  
tatttagggt gcacaggcac aagtagaaga ggcttgcac atgcggaaga tacttgacaa 46800  
tgaagattcc tctcagaaaa catccactgg aaagtccggt gcagtaattg ctgccagagt 46860  
tgcagataag cttgcatggt cctgcgcgaa aagattgggt atcgtaaaaa ccattcaaaa 46920  
ttaagatcgt gtatgtttgg tttgagttgt ctaaaataa atgatgtgaa agttttttaa 46980  
aaaatcgatg gaaacattga cttgggttcag aattttcttt tctatttgag agccaaaacta 47040  
aatacgacta atttaatttg aaataaaaag aatcaattaa tcaaatattt tgatgaaccc 47100  
ctctataaag aataactttc tgtgattatt actatttaag tcataagatg acggggggga 47160  
taaaagtac tcgcagtatt ttataaatgg gctgaaactc caacatcaac tagactgtcg 47220  
atacagtttc atactaaact aaatctgtac atatatgtc atcgttatcc agtttgagaa 47280  
ttgaaggatt gcaatttgca aagaggtcat aatccaccct cgaagtctcg gaaagaggaa 47340  
gcaagcaata agcagaactg cacaactctg cattgtgcat tactgttttg gaacagtggc 47400  
aggaactgta ccagatttgt gcttttaata actgaagggg acatgaggct gtctgtctct 47460  
gtcacgctcc aaattagatg tcaagttaag aagtttttga ggatgtaatt caattccatt 47520  
cctatcataa gagecgtgta agtcaaactt atgaggtggg caaccataca aacatgaaaa 47580  
aacacattta taatgttgag aaaaactaac aaatgcagaa aaaaagtata gctggtctat 47640  
gttcaagaa tgaggaaagt tcagacacat tctcaagaaa ctatgtcttc agtttggtca 47700  
agtttttatt aaaaactttg ttactataaa tatgataaaa tggaactaaa tagaaaaat 47760  
ctcttctggt gctaaaggta aagtaaatgc ttataagggt gtccatgtgg aaagatagat 47820  
gcaccacctc actatagctc agcctttcat tgctttcttc aaaacaaaa ctctaacaga 47880  
gatcggtaaa tgatacatga tagccatgac agtagagtaa tgcccatagg tgaaccatgc 47940  
aggtggttcc tccttaagga cagcagctac agtgcttata gcaaacctat ctgtagggg 48000

tgttttggac ttctgcgaaa aataagctct atctcggatt gctgcttcaa aaggcttgaa 48060  
 taatttccat tctggcatgc ggttgtagct ggctatggca gaatctccaa tgtttgattt 48120  
 gatagctccg gggacaatat tgacaacatc aattccaaaa tgtccaagtt ccagttctgta 48180  
 taccacatgc tctgtttaa gatgagaaca attagataaa atgtaaaata ttggcacaaa 48240  
 ggcagccac cgaatcccta attgttgcat ttgcatccac tagattttga tttgaagtg 48300  
 catatataga gttttacaaa ctccactata ttgcagtttc catgaaatca acacattgat 48360  
 ctggaccaac caaagcctaa aagcttgctc ataccaaacc aacaatttct atagaacaaa 48420  
 ggtggatttc tgtagccaat tactgttgcc tctagtgtat ttgatgcaga aatgaagcac 48480  
 aggtgtagag taacgcaagt ttgaagaaga aaaataacct aatgtatctg tcaaagcatg 48540  
 aagagcagct ttgaagctg tataagctcc tgaccaaggt ccagagacca aggcagcaac 48600  
 actaccaata ttgacaatct tccccttttt cttagtgtcc atatgaggaa caaccgcctg 48660  
 aaccattctc aaggaaacct catgcgtatg cccgatgaaa catagcatca agtttaacaa 48720  
 gcaattcatg gttttgtaca ccaaggcaat ttttgataa aaattgtgaa tgtgattttg 48780  
 gaaaatccaa tgtacattct tgaatttcaa cacacaaaat aaagaaataa gaatttgtgc 48840  
 atgcttgaat gaaagttaa aatattgtgt gtgcgtttca aaaaacagat aaaaataaaa 48900  
 tagaacaaca agcaaaagtc ctgacaaaat tacttttttg tctcaaataa tttttaacct 48960  
 aaaaatcaaa cattatagtt tgaggagtaa aagtactttt tgagttccaa aacatgtact 49020  
 aattcagttt gcgcataacc tttgaggcca caattgacaa ataagaatga aaacaggcaa 49080  
 cagcaactta aatagagtgg ctggttccaa aatgcattat ttatctagct aggagtaact 49140  
 gtagtgaaca gaagccatct gcagaatgat taaaatcaat ctaaaacaaa ctcatattggt 49200  
 taatctaato caacatggca gattaacaac aaatcccatc caacaatatg catcaaaaatt 49260  
 ccaattaaga taagaatac tgaccgaaga cattgtgtgc gaaagtgttt tggatggcag 49320  
 agagaggagc ctgcgcgagt gggccacac actgaacacc gcggttgta acgagcacgt 49380  
 cgatcgacc gtacttgtcg acaacagcat caaccacctc acgcacgctc tcatcgact 49440  
 gaacatccaa ctcttccaaa aagaaccttt ggtcgtgttc cagctccgcc atgctcgacc 49500  
 gcgacctgct ggtggccacc accctgcact tcttctcagc gaaggcgcgt gcaagcgcgt 49560  
 ggccatcccc tcccggtgaa catcctgtga tcagaaccac aggtttcggt ggctcctcct 49620  
 catcgtggcc gcgcgtgtcc atcgtagaat acaagagag gcagaggagg tgatttagatt 49680  
 agtgtaact tgtgtctatg ccaacaatga caaccaccc caaccatct agagtacaaa 49740

acaaaggcaa gacttttatg ccagcttttc ctatatctag agtagtactc ctacaatcaa 49800  
atgacaagca gctacaacca attttctctg toattcattt attacggctt atacatttat 49860  
catttatgac aactattaat taataacatt atttttcacg gtcaaaaaata agaaaaagaa 49920  
aaactttatt ttgagttta tattagatga taatatttat gtttttttta gaaaaaaat 49980  
gtctttttta tttttatttt tggatcacac tactatatcc attcccttac ccacatctaa 50040  
aactcgggtg gaacggtagc tcgagtgcac gttggtttgt catgcttatg gacaaactaat 50100  
cttatattct cttttaactg aaatagtagc ttaagaaata aaatactact ttttttggtta 50160  
aatagagaaa ataagactca gaagacgtca cacttaaaaa tcttcttaac aggtttcaat 50220  
ttattcaagt ttgtaaaata gattttcaat acttcagaaa tgtttgtctc atctcgatcc 50280  
gttaagaccg taatttgggt agattatgtt atttctttac tgaagcacg gtaaatattg 50340  
aaggcaactg tgactacatt actactatag aaaaggtaga aagtatgttt gataaccgag 50400  
gtttcaaaac aaaattaaca caatctcctg aaatgacatt aagcctactg gctaccggaa 50460  
agcaatctac tcagcagcat cttttttctt cttgggcatt ctgtatccac caaccacaca 50520  
ggcatggtcc ctctcaaatg gttcaagggg gacttgctcg aatggtttga attgatctgc 50580  
cttcaacttg ttactctgc tttcaaacac tgcctcgga gggactgtgg agtctatgca 50640  
gttagcctaa acatacaaat gaattcaagt gagaacaatt tagcagatag ttaagacata 50700  
gtaatacata agagaagaaa ctaaacactt taattcaagg taatcataaa catttctctt 50760  
tccaatgttt tcaatagagg gcaagcctcg acactccat atgtttgttg tttggaggcc 50820  
tgaaagtccc aagttcaaat tctaaaaaca gccttagagg gtaaaagcct gattccttct 50880  
actctctta gacctacta ggtggggacc atgtgcaact ggtcacctt ttttaaaatg 50940  
tttaccata tgattactca ataaagtttt ataaaaacat gattagatgt ctgaacatc 51000  
cttgattgaa atacaacaaat gacctctgc tttcagataa tatgaagcat tcagcccaaa 51060  
aatccttgcc tgcaatggtc accaaaaatc ataaacgaaa ctgatacata gaatataact 51120  
ggtgtttcaa aaatagtgtt ttgataatgt ttaaggagc ttttatcttc aaatagsata 51180  
taattcattt ccagaaaatg taaaaaggaa tocacaaata taaaacctat gcaccactcc 51240  
aaccacagcc acactttctt cttaaacagg ctgcagatat aaactatcat aacacatcat 51300  
cacactctct ggaccataat ccaattgagg ctgcttataa tatctgacgg cccaccccaa 51360  
tacctgtaac accctgctaa tttgcagccc ctgaccagc cgggggattt aaaaaaaag 51420  
acagtgtata aacaaaatcc tttctctttg ccttcatcca ctccaagaat tgaagttaac 51480



catctccact atatcacatt ctcatacata aattgaaaa aaaaaagta cttatagatt 51540  
 caaaatgggg ccataaacag gctaagaaaa gataatccag cactgctgtc attttaaaaa 51600  
 attaaagcaa taaaaataa gcaaaaagga aaagaaaatt atgcgcatca gaaactcggg 51660  
 tctgcatcat ccatgatgtc ttgaagggga atccgagtaa ctcatcatat gcacagatta 51720  
 ttccaattta acagaataca tgcaagtaaa aaaataatat tgaaaattgt aaaatacaaa 51780  
 tttgagaaac taaaagaaaa acaacaaaat gaactctaaa ctgagatttt aaaaattctt 51840  
 tcattttgaa aatcgaaagt tatacacaaa tgtaaaagag aaatgcaac aaaaataatac 51900  
 acgtgtagat ggtctggctaa tatttggcaa ccaaacgaaa tagtagatga attgaaatta 51960  
 aacaacaat aatcatatgg cactgcccgt taagtacatc agtattatgg aaaggtagta 52020  
 tgtattgtct gtaaaaaaga gactgaaacc aaactcgaga tatgttgtaa gaaatataaa 52080  
 ccaaactga tcgggctgag caacatcaga aaatatgaca tcaaccatc cgaccagcat 52140  
 cctatactta gctggatgtc tagcatcttc aataatggga ataacattag ttcgcttctt 52200  
 tgccatatta accaggctac gtccactctt atgggaaaat tctactgcat agacaactcc 52260  
 agtcttcata caagaacaat aacacattat aatcagcact tctattccac aaactaaga 52320  
 ttatgcaatt ctaaatctta aattttactt acagggccaa caacgtctga tacatgagaa 52380  
 acagtgttcc cagaagcagc tcctaggtaa aggaactctg ctccaggttt cttcataaga 52440  
 cacaagacaa ataatttaa aaaatgaagt gtgatttgaa gaacttcgcc taacatagc 52500  
 acatgttgta aagacacat cttacaatcc atatgttgtc aactccacca aggatggcag 52560  
 cagccaactt ggatcgaaaa ggattccaaa ctctgtactc atctttggaa ccatcttctg 52620  
 tctgcaaaat gaaactatca attcatgcat ccatggttca gtacttaaga agaaaaaca 52680  
 aataatggtg agaagaaaa gaaacctgaa cagtaactct tttctcattg taacagcctt 52740  
 caccaggaac aagattctta gtaacaagag cgtcttcttt acccttggca atgaaaaatac 52800  
 catcatgtct atggggctga accacaacct tgcctcctcc ctctattcca cctctgccac 52860  
 caccacggcc accaccctg cctccgccac ggccaccacc acctctgcca cctcctctag 52920  
 ctttgaaggg agtgccctgt tctcctctcc ttcctccacc gccctctctc ctgcctctgt 52980  
 caccctacc accaccctg aaccacccgc caccacaaa accaccacga cctgaacaaa 53040  
 ttcaaacact tgtaacaac aaaaaaaaa aaactattac cacaacatc ttacgtcacc 53100  
 aaagaccata tcatttagtt caaatgccaa atgccaaatg caaaaaccaa aatcatcaaa 53160  
 ccacattttc ttattaaaaa gaaaattacc tcttgagga gccatttttg ggagctgtaa 53220

actgaagtgc aacgcagcag caacaactgg ttaagaggat tttatagcca gccgaacaac 53280  
tgaaatagtt aggggttaat agagggtagg gttttgattt tatagaaaa gaaaagaaa 53340  
gtgtcttgca tcatcacta acctaacctt agcggtcggc ccattttcct taccctgctc 53400  
acttggtcgt gggcctttat ttgggttcct cagtcttcct aacaagcatt tattgctttg 53460  
ggccactagc aaggatttct ctttccttcc tcagctctct acggattaat taagcttttg 53520  
atatcttcaa caaaaaataa ttaattaaaa tcaatgaact cataatattt ttttaaaaga 53580  
actaaggagg gtatcaagaa tatgatgttc ttcttgatat attcagcaac aattttagaa 53640  
tttgctcgag attgatgttc ttctttactc aacctcttat ttgactccat gatttgttcc 53700  
tcatacaaat aaattatttc ttcatattgt gtaaccttca attcaactcc tccattttgtg 53760  
acaacaaatg attcacatcc acattacttc cacaaaccac atattacccc gccaccatta 53820  
ctttcaaatg ccttgtttaa tgatttcttc acatcgctat cccaaaccac aaatctatag 53880  
ttcttaactt tctacaatca caaacaaca aaagcaattt aaaaccatag aatagaacac 53940  
aacaccaat aaagagcata aatccttaca ttccaccaag cacatcacca aaataacttg 54000  
tttgatgtct agttgggttg tgatattaat taaattgttg gcttctcata gagacaatga 54060  
acaacaccac caaatccaat tgacgaagac aatgctccaa atgatgccat ccacgatgaa 54120  
cacgaacacc atgaacccta cccgaccctc ataaaaccac taagagccac ccctaactca 54180  
aatttgctct ctttgacaat aaaacatag attatgctac tctactcgtc ctagtctcaa 54240  
gtagttcacc taaaaacca tagataatct aaaacaaaac gataacattt acacaattta 54300  
ctatatgatg cgtatgggca tgtccaccga gacgaaccac atggatctcc cacaaaactt 54360  
ttgataagaa ttcatgtaga aataaaaaaa tttcaaggat cggtcacgga attcgatat 54420  
attttaagga tctccaacat gtttaagtct tgcataatta ttttgattgt tataaaagga 54480  
aaagggaaaa atatttaaat tcttaagaaa ataagtgtatt atttttattt aattttttat 54540  
atctctttat acagttttgt ttaaaagata aaataggaaa actactttca actctctcct 54600  
agataacca atattatact attctccctt ctccaattta aattaccaat catgtatgga 54660  
gtgagaaaac ctcgagcata tgagaaaggg gtatatgaat ataactactt tcgtattatt 54720  
atcaaatatt aatattgttc atagaaattt acaaatattt tatattctat tattattatt 54780  
tagtaatagt agtatagat tagtaatgac attggatcgt aaatgtaaa taatatattt 54840  
atcattagta tttgtaggaa tgtaaaattt aatttaaatg aaaaatatta tatgacaaaa 54900  
gttatgtgac aatattgtgc ctggttgtag tcagccattt ttttgtgtgg cttataacat 54960

acctttctaa tagatttttt ttttattgac tgaattttat taaaaaata tattattttct 55020  
 ctccaatttt atatactttt taaattaata ttaactaata ataaatcgaa cacgtgttaa 55080  
 atagtgtttt gtttgccttt ctattgttaa aaatttactt ttttattggt aaaaaatac 55140  
 tactaaattt accttcattc taattcaaat taattaaaa ttatgcctta tagcatatta 55200  
 tataaaaaat ttaaaaaatt aaatgtttca atttttttaa aatagtttca aatcaagtct 55260  
 tactcaaat aatgtttttt ccatctttaa aagatgcata atgcagtttt aatttaata 55320  
 aaactgaaga atgcattgaga gctaaggatg cattattcat acgtatgaaa tgacagaaag 55380  
 tcttatctgt tttcttggtt atgagaaaaa tatgacgaaa cgaccttagt gagtaatttt 55440  
 tttaaaatta taaaatttaa ataatactta aaaaaaata agttatagaa tatttaataa 55500  
 tttctgagtt cataaattat ttgataactt ttgttttttt ttattgaaga tataaaaaat 55560  
 tatgattttt taattttttt aaaattcaat tttaaaattc taaataattt tattttaatt 55620  
 aatattttta ttttaatttc attttatatt tttatatatt attttatatta atgtttttat 55680  
 atcattttat ttaataataa taatataaaa atattaaata aaaaatagac aaaaatatta 55740  
 aataaaacta aaaaataaaa tataggactt tgtatattgt tttttttcag tgagtatttc 55800  
 ttacagtagt ttactgagga aataattagt taaagaaatg ggttttgggc attttgggag 55860  
 gcaagcaagc ccatcttcca acaagcaaat tgtttcccc acagcagcac ttggcacaata 55920  
 gtgattgatt aatcgctgtg gttgtggctg tgaccgtcga aggcagcgac ggagtgaagc 55980  
 tccgtacaag ggtgttcaag ccacagaaga agcagatggg aacttgggga ttgttctgct 56040  
 gcaccogtat tctatttttg gtggctgcca ggctctcttc aatggttaca ccgctttcac 56100  
 ctttgacatg agaggogttg ccaagttcac cggcttcttc caagttaagg atgttgttgc 56160  
 tgtctgcaac tgcctttcca acactttctc tcttcccgag attttgttgc ttggttcttc 56220  
 cgcaggtagt ctctattctc agatctctca cttttcttat actgtttaac atcaccccat 56280  
 ccgctttttc tacctacaaa tcttacaggt atcattaaat ttctagagta ttctttttaa 56340  
 tttttactgt gaataatgt gtcttttttt ttttttttaa taatctttga actgggcttg 56400  
 ttgtttcagt atatcaaat tctgaattgc ttgattagct cctagctgta tgtttcataa 56460  
 acaaatctg agctttctgt ctccctgatt acattttggt tattccggcc agccttatgt 56520  
 caaatttga aaggatgcct tagttacttg ttgcattgct gataagtta gtttgagagt 56580  
 ttaaatctta atgccttcca agtttgaaag ttgtttgatg ttgggataaa cactccatgt 56640  
 gaattaaaa atgcaatgac agctctatcg actctcagtc tcaggtttta agcttgtagg 56700

agttgaaagg aagtatatat cttttgtctt tggattcgat atagactgat ttgcagatca 56760  
 acttgtttcc cattttaagt gttggaatga tatccaaacta attatgcattg cttgagttta 56820  
 catggactgt tagtttagact tgtggtgta atccacatgc ttcttgtctg tgacaatttc 56880  
 cccagttca ttcacgagct gctttaggag atgtagtaat tcgcgccatg tggttgatca 56940  
 atgttgaagc ctgaaagagt tgtcatatat ttgagttatt tcttttatta cagggtgcacc 57000  
 tatagcaggg tctgccctgt atcagattga acaagtcatt gggtatgtaa gcatagggtta 57060  
 tccattggga atgactgctt caatcctttt tggatgacac cataaagcca tctacagtc 57120  
 accaaaacca cttcattatg ggaacacaag atggattcac cagtgtaaag cagctgagga 57180  
 acaagcttaa ttctcgagca ggacgtgtag aaacacattt aatcgatggg gtaggcatt 57240  
 tccagatgga agggcctggt tatgatgctg agatgggtgga tctcatcacc aagtttattg 57300  
 catcactata aagacaaata actgattgtg ctgagtgta ttgactctg ctattgggaa 57360  
 ttgggattag tttgtgtaca aaatcttctc agaatctgct atttgcactg tttatttctg 57420  
 tcccctaagc ccatgccacc ttctaaaatg aaagaaatca gttaccaca tagtaacat 57480  
 ttctgctgat tggataatga gagagaatgg gaattgcaag cataagaacc atgcttttca 57540  
 taatttataa tttacaaata acatatctgt tgtctctcta gattgaagct accctagatt 57600  
 gaagctacc cctacactac ttaatagata caaaaagaaa ctacaattat acaataatta 57660  
 tatcatatta caatagtcac gcttgttctg caactctgtg cggatactca catggagcac 57720  
 tagtttctct tggtaattta tcaccatttg aaccgtctgg tactctaaca ccaactcgaat 57780  
 tatactgttg tttcttccag tgttttccct gatccattga ctggtcttgc gaaaaggag 57840  
 gttctacatc tggctaaagc aagagagttc aaggaacata aagaatgaaa gagtcttcta 57900  
 atcaataatc taaaatatca aggaagttat atactagcaa aagaacaaa aaaaagacgt 57960  
 ccaaagctcc acgctttatt cacagccttg catttcataa aaaagttcat tagatcccc 58020  
 gagttgctgg cctactagca atatccaaaa attagatgc aattcaacag aatttgggtga 58080  
 aatttcaatt aattatacat gaacaatttt ctagttaact agttattacc actttcacc 58140  
 tcagacacat ctcogagcct ggtaattgca ttaacaagag cctgttcatc ttctctgatt 58200  
 tcaaagggag agccatcatc attttatttt atttttgat atagagagcc atcattattg 58260  
 ttgaaactaa agatttttca taacataaag ttgtatacta agaaatgat tgatattcac 58320  
 ttttagcatt ttcttctgtt tctccatttc catgggatca ggatggcttg cagtaaacac 58380  
 cttttcaacc tgttttgggt ccattctatt aggtaatatg aaggagccaa atatttgtat 58440

catacagaga agtagccaaa tttaacacac aaatttgga atttctgttc cagatgcttg 58500  
 tacctccttg ataagggtat ctgatatggag tatctcaata tcattttagt ccttcttctt 58560  
 gccctctttt gtgacaaaagt gaaatctttc ttaggctgca ccttcacatc tcccctacct 58620  
 cttccagaac cagtaactgc accaccgtga ggcataagact tcttcaatcc tctacctgga 58680  
 cggggccggc cacctttacc agatattcca ggatcttgac cctcccaacg aatatcttct 58740  
 ggagatatct accatatcga atagatagta actaggattt catcactgat tgagttgatt 58800  
 gttagcagca acaacacaaa ctttagtcta attgataaag aataatgaat gtgaagagac 58860  
 actaagcatg taactatgta tctttgaagt agagggtggg tctgattttt aaagacaatc 58920  
 atcccgata tgcttgaagt ttcattgttt gcattaatta tagatacaca aaaaagaat 58980  
 caccaaaagca aatgtttgatc caagatagca aaaggacatt gactatatgc caatcccttc 59040  
 agccacaaaa acatactaga ttttagtttt aggcagaaaa atgttcattt catcaaaatc 59100  
 atgagatact taatacccat ttcaagggtat gggattcaaa gtatgagaca caaggtcacc 59160  
 atgggaaagg taagatatgt gaaaacaaaa ggttctagct atgcttgta ggaagagaat 59220  
 gttgcttaaa aagaaaagaa aagaatgaac gacgctaaaa ctgatttgta tgtgaatgat 59280  
 agacagaaga atattaatat tgtaaaattg acctctttaa gattgaccca ctcccatggt 59340  
 tcatttgctg aattcatatc ataaacaaa gcatgtagtc cctacaaaat atgtaaatag 59400  
 aaatattaaa ggagtcctatg cattcctcag tttctaatag tatactataa acacacctca 59460  
 gcagcattgt tgtttaggga acaagcttat gagactggat tcgaggccag cacctccaga 59520  
 aggataagaa aaggcaagtg taagaaggaa gcttgataat tttcatatat tccatttgaa 59580  
 ttacaattgc tttatttata ctagtctcgt gcaaacagaa tttgtcattt tgtaccaaga 59640  
 aatatttagg aatcattgg cctagtgatt acggttgta tgctgtcccc caagaccaa 59700  
 ccagccacgc ttagcaattc cagagcaaat tcttctatc gaacctctt catacgtagt 59760  
 cactaagggt agagggcctt gtgatgacct tcttgcctt tcttgctatc tgaacccctc 59820  
 tgttctgtgc ctgcacctct gtttgatgtt ctttgctagc tgatgatcct cttttcgtat 59880  
 caattataat ccgttataac agcttcatag aagtgggtgt cttcaggcca ccttgccaa 59940  
 acttctcttc caatccaggg atcataagaa gctccttcag caggtcctt tgcgtcaaa 60000  
 gcacctgaag aacccctga tcttgcttt ggcttgtaag tctgtgaa gggaaactgta 60060  
 tggacatcat tggagatgca caaatcaatg gggtttttct ttagaacggg atagaatgac 60120  
 catagaggaa gtgtaactca gtattttgtt atcttgatat tccaaaacaa aaaggattac 60180

tggatTTTTgc tttgtggtct tagttcctga tggagggtca tgcttcaatg cggatgatga 60240  
 tggttgcata gatggatgaa ctgcaggaga tgggtgcaccc aaggataatg aagccacaga 60300  
 ttgtgatgtc tttttttctt tacgagatgc caaaacggta gagctgggat ttgtgatcatg 60360  
 aacctgttga gcagtagtga ctgttccaga ttggagccca ttccctttc tccactccct 60420  
 aacagcacia ttggatacac cccactttaa aatgcagaaa aaacctgtt ttgtctaatg 60480  
 ctatacagaa actgctcact agaatatgat aaaatatagt atgaccttat cctgtggatg 60540  
 atgtcatcag cattgactct tgateagaagt tctgtgttct tcatctgata ctcttagctc 60600  
 cttcctaagc tctgttaatta aactttcttt ttccataagt acataggaac aaaaaaacctc 60660  
 attgccaatc ctatcgacat acttaataga aagacaaacc aatgaattaa ggcaagggtg 60720  
 cccaagtgat tgcacagat tgagctttaa aagccttcag gacagagggt tatgtctctt 60780  
 gctcaatggt atggatttgt gtttccatgt cgttctgcat ccgaggatat ggaccagagc 60840  
 ctacaattgc agatcttcgc tttccagttc cagtggcatg atcagggcgt tgaatctat 60900  
 ttcgacattg atgggaagga ggaagatcat catctgttcc ttgtggtatt tcagattatg 60960  
 gaaattaatt gcaagcaacc ttttgaacg ttactcata taatttccac atctgaaatg 61020  
 ttactcatc attctatgag atttatcttt atcaaagaaa atttaggaca gtcggacaga 61080  
 tcaatcatc ggtctgtact taacattgtg cacaacaatt atgcctgtgt gatagatatc 61140  
 ttgtgttcta taacgatgat aataattaga aaggagcgc atagatatatt tgtacgaaaa 61200  
 aaattgtctc ctgattgat ttgtgtgaat ttggtattgg ggagctggta gtggtactta 61260  
 ttttgattat ttatgggtatt cccctgctt caaatgctt acaaaaaag acaaacctga 61320  
 gaatatgctc atttcagtta tgaatcagac atgactcatt tccacttttt ggcaatatat 61380  
 tctgttcttt tcattctgat ttacagcaa gtgttcatta tcaatgacat ggtctgcgtt 61440  
 tggactaaga tgacagatca cgaggaaaa tgcagctagt tcatgtcttt ttttcttgc 61500  
 attttctga cttgttctca gttagaaccg ttttcttga aacactataa ctgctccatt 61560  
 atttgaatt aaaaagtaata agaaaattac aaaatcagcc ttgactgtgt tcatttctag 61620  
 gatttaggat ataataagt gacaccaaag ttcaagtctt ttaagtgaag acctcaatct 61680  
 atgcataata aaaaaattag gattagcata aaggagaggt caaacgaagt tattcaagtc 61740  
 agttgtcaca cacctgcata attttgtaac catacatata atataagcct gcgaagaatc 61800  
 accagtcac tatcaaaaaa ttccctaact gaagttaact aagcaatttg ataagaaaaa 61860  
 ccaaaatcct aatttccaaa ttaagggtca aactcattct agtcattatc tattcaaaaa 61920

ccttctccaa catcacaaag cagctacaat tattctcaat attgagtaa catctcaaat 61980  
 gaaaagtcca caaagtgtga gtttcaggca acaaaagtcc aacttacaac aataacagca 62040  
 gcagccaaac ggtttgttga ggaataagaa tagcaaaagg gctatttcag acaatttcct 62100  
 caattttgtt cacacggcca ttttaatat ttgaattata atttattaga ggccgggtct 62160  
 agccaaaaat attagtcctc ggtacatccg catttaaac ttgaataaga ttgtcttttg 62220  
 aagaataata cttataagac actcaacgaa attattttaa atcatttgat ttattattata 62280  
 tgtattttat aattttacac aaatagtgc aactaatgta taccacaaga tctgcttgga 62340  
 aaaagttaaa tgtatactca attttatttt gtatgcatag agagaaatag acgaggaaaa 62400  
 atataaaaga aagaaaaaaa attataata aaacaacaaa cataataaaa aatgaaaaga 62460  
 agaatagaag gaataataga ctcaaatga gattttgaaa ttatataatt aataatccaa 62520  
 ttaatctcat gagatattat tctataacta taacttatgg accgtttgga agagataaaa 62580  
 aaagaaaaga aaatattgat tacaatatat tttcttttgt ttaattgaag aaaaaaag 62640  
 aaacaattga cttctaagaa taaaaattca aaacttttct tatgttttaa ctttttttta 62700  
 atttcaaat ttcttttct ttgatttctt ttcccactac aattgactga gtcagactga 62760  
 acgtgatact tagatacata catgcaactc tcagtactac tcctcacctc tatataatct 62820  
 tcaactcttg ctctctctt tcttcatccc tcagaattaa cctcttctct tctttgtgccc 62880  
 gaggtattgc tctctgtctc ccccatcttt ttctaataa atgcactatt actccatata 62940  
 tttgtctctc atctctcccc actctttaac taataagatt ttctctccgt tgtgtatgtg 63000  
 tatgtgtgca taactcgatc tcttgatgat ttctaccaat tggcctcggc cttatatata 63060  
 tatattctca ttctcatata tatatatatg tgcgtgtagt ttacatgaat ttgacaggtg 63120  
 attattattt gaaagcaaat ttattttatg tgtagtaaa gattgcagca agattttgcc 63180  
 ataattaaaa atctacttta accccattgc cctgcattga cattttatta ttagtttgtt 63240  
 tgatgatttt ggaattttgt ggtttaattt attaatatat atatatatat atatatatat 63300  
 atatatatat atgtatatat tttgaaggaa ttgaggagcg gagtatata gatgcagttc 63360  
 ggattaagat cccagaagt accagatgag aacacaatct cgcgtggaag catcactata 63420  
 cagtttagct tccatctgag tgcactgagc atgatggagt atgagaacca tcggagagtg 63480  
 agaacactga gcacaggctc cttcacactg aaaatgataa tgcaataaaa ctgtcccat 63540  
 cttgtcata tctgttgtgt agataaattt aagtggcctc taacaaggaa cactgttgtt 63600  
 ctccgagcca ccactgcgac ctatgtcatt gccctcctg gcccttata tgcctccag 63660

ttccaagac tttgtctctg agaacgtctt cgaatcctct caaatgtctt taggaatat	63720
ggctacttcc aggatcttag accccaaaag ctagggtatct ctcttctcat taatttcaac	63780
ttttatttaa ttttcgttac cattgacttt tattgattac tatttgatta ttgtattgac	63840
tacttttaat aacattatct cttttgacta cagggtatat ctatccataa cgettgactg	63900
ttacttagac cttcaacttt ggattaaaat tttaaaaata atttttgaat ttttaaaaaa	63960
tctaatttat taataagaac tgcacacat ttttttgcgt gaatgcaagg ggcaaaagcc	64020
cctagaacta gattaatacc ctctcgggaa aaccgtggaa ttacattag atataacagc	64080
tgcagcaaag aaaggtaata gaatcataaa ctctcggatt ctcaacaaga gacatttcat	64140
ggtagctaa ctctcagta cctgattccc ttccctcaag acaccatgat tgaaggcact	64200
ggactatact aaagaactgc tacacatgcg tatttttatt tttttaaat acctaattac	64260
aaataaattt gtactaatth tcttcccttt tgcgattggg tgcaaaatta aatttattta	64320
gacatacgat cgaccatttg tttgtctac atctattat ttattttgag aggatatgtc	64380
cttgtcctaa ttatataat gaatgaatga tgacattgat tataattaat aattaattac	64440
attacatgca ggtgggtgct caatgcagtt cgaatgacca gaaatgttg aagtgttacg	64500
tcctgatttt gagcccatag cacatgaac ccttttgagg ttagggaata tccagggat	64560
ttcccgtttc tgcattggaag agcagaactc gagcttcta agggcgtgtc gcatgtcggc	64620
gatgacaaag atgataatgg aaagcatgag tgtgaagaca atgacaagta gtcttgtgga	64680
gatcaaaagg accaaccacc ctgaggcaca ccaaggaaga gcctatgctg tgatgagtac	64740
ctacactcgc cttgttgatt ctgctgaata cgettccctc gtttccctgt gcttagttga	64800
ggaattggag aatgtttgtt ctctctctcg ctttgggtcg gttgatgttg ggtaccaggt	64860
ttggaatctc aatcacttgg gccttttgc tttcatgtct ctcttgggac cccatgttga	64920
agaggaagtg aaatatgtga gtgctagaga agagtgttg aatgccatgg ctgaaactga	64980
acaggaagaa gaagctgatg aggcctctga tcatgatgaa gataagggcg aggaaacaaa	65040
ttaaatatca atacataagt ctggctagat catgtatgag gtcttgagg ctacattta	65100
tgttcgtgta ataatttctc ttttaacaaga gatcgagggg aaacatatg tcttaattaa	65160
ttttctattt tgatattatt atatatctc tactattata atatgatgaa gaactaaac	65220
atatgaattg attgattttt taaagagtac aatcgatat atatggagag agaattggga	65280
agtgaccata tgggatgacc atcacctagt aataataagg caagctaaag aaatattatc	65340
atgataactg ttgggaaaat ggaatcggtc ggaagcatgc atgtggaaca gatgaaatta	65400



aagctagctt ctcaggaata atatgaaaat ctatgtcaat atgctagctt ttcgcaatca 65460  
tgaaaattaa ttagtattgg atatatattc gatattgtac aatacaataa tgcaggagtg 65520  
aagaatggaa taaataatct tgtagaaggt aagtcagtc tgaaaactca catgtggtag 65580  
agagggctga aaggctaaag cagcatatta tatacagcat ttcagacgag gagcaagaca 65640  
gtacatatat aatagtatct tgtttcttga atttccagaa aatcaaagag ccacgcagga 65700  
acacacagaa tccgggtgtg tatgactaga tgccaaattt gaattactaa atgcacgttt 65760  
agtattgaag cgaatatctc tcttagctgc ccgaaaatga tgaatgggtg gaacctgggt 65820  
gaaatggcta agatgctaaa catcatatag gttattatat ctggctctta gtttgtcaaa 65880  
tgaagaagac gtccaatata ttttaacttt ccagctagc tgcgttaaca tatattgttt 65940  
ttctcgatat atcttgaaag gtttttaaaa attcaaagtg attatttcta taaaaaaaa 66000  
gtttctaaaa ctacaaaaat gttgtataca ttaataatat taataatttc agtggttagat 66060  
tgattaaagt tcgctgatat aacttttatt taatgaaaat taaattgtta gtacaaatat 66120  
tgcaagctct tattaataat acaatacatc atactaacat atccaaaaa tttatactaa 66180  
ttaacagttc aatcctgcta gataaatgct ttctatatata tactactgaca gaacagtttt 66240  
aatgcctttt taattttttc ctggttgtaa taatgatatt ttgtttgttt gttagaaatt 66300  
gacatttttt tcactttaat catgggctaa atcttttata gtaacaaaag aatataaaaa 66360  
taaggaagtt taacttagtt tgtttcaaca atataagtaa attattgtta attctcttat 66420  
atctgaattc ctacatatta aaaagaatat tacaaaaata agatgtatat taatttgatt 66480  
tgataaaaaa aagattatgc aagatttaat taataaagca aaataataaa aacaattaaag 66540  
tcttttaaaa ctatataaag tcgtctattt gaactaaagt tatcataaca ttttatgggtg 66600  
aatcatgtgc ataataaat tattgatctt ttaataatta tttttaatgg attcatcaat 66660  
agtttttttc agtatccttc tacctctagt aatatggcca tcacatcatc ccttttgatc 66720  
tactctctta attggtattt tactagtttt tacatagatg tcatgtctac ataactttat 66780  
ctttactatg atttttctaa caatatatta tcttgatatat tcactcattt agtgcaccca 66840  
tatttgtaat atttggtttt attaatgttc aagacactcc tttttgtttt tttaaagtta 66900  
attaagatac tcattcattt catctatcat gcttaaatcc tctagttaat ttacatttct 66960  
ttctcattct atttgattga tggatccaat acttcagcat gatgcacgtt attattgggtg 67020  
acaacttata atagtcatac tactattaac aaaagataaa cttaacgcct tgtgcaaaag 67080  
aagattacta tcttaggtgt agtgtatatt ttgtgttgtt ttgtcattcg gaagataata 67140

ataacttatt ttatcttac ttgcgacaa ttcaaatct ggtatgtaat ttgtgtatgt 67200  
 caggatattt gagatgttg ataataatat ttccgccac tatatacaat ttaactctct 67260  
 ttttggaaga aaaaaagtg aggattaatt gcaacaacta tatggtgatc cttttaactt 67320  
 caatacgttt tcgacttgat cgataaaagt tagatcataa acataattta aggtcatgtt 67380  
 cttttttcta atggtgtatt gatattctta ttatccctt taattttact ttaattgtaa 67440  
 gttagacttt actttcatth taataaaatt ttactattaa aaaaataaa acttatccct 67500  
 taataagtga catttgtttt tgtacaataa ttaatgtgtc ttattttaag tgggtggttag 67560  
 atccaatttc atgttgaatt taatattgtg gaacaaaaag gcggtgttat atatataat 67620  
 atgatggtcg atggtgcaag aggggatcgg aagattatac ttaagaagc agaagcatac 67680  
 ttgtactcca atctaataaa aatgttctat atataataga aaatacggat acgtgctgta 67740  
 caggaaaaga aaatgaatat gtaagtatga ttccgaatac atatgcacca tgacgctttt 67800  
 aatcatggat acaaacgacg aaagccctct ttgacataa caacacattc aattttgaga 67860  
 catgcatgtt acaaacgact aaagccac gccattaata acatcaacag tcttttatca 67920  
 cttctatata gtaactaaac aaggtttttt ttccgtgagc ttaggggagc tagtcaaa 67980  
 atcatttaaa acttaatac acaaacatg tagaggata gaaaaatac aaagggtaat 68040  
 ggtgggcaag tgaataacca tagttttggg gaggttattg tatccaaag tcgaaaacca 68100  
 ttatgccta aaggcattat cggatgtgac tcttgttca ctacctggt gcacctgtt 68160  
 ctttccaca ctagaactag ccacctttct ctcttcagta ctacctca ctctctgtac 68220  
 tctttccact tccattctct tcaactctc ttctgcctc ttctccataa cataaatttc 68280  
 agctgaacgt gcccacatgg aagacatgtt gttctctgca aaaaacaaa actagcta 68340  
 attgcaaaa agttactaa ttattagact tggctactaa gaaatgcttt atgaatttag 68400  
 cattgcaaa gatggtattt ataggagttg gaatttgatg agagcgtgtg gaataagcag 68460  
 ccacgttcat ttctcaaat ccctgttgtt ttgtggcttt gctttacaac caaccataat 68520  
 taacattaat tatcttaatt gcactcacac aaaaaataa aatgccatgt tgtcatcata 68580  
 ctttgcatgg aaaattaagt ttgtgtccg gcacatctt tatgtgtttt aaactttcta 68640  
 aatatataag atagacaatc ctatggctcc aatgcaaca attgtatta atatattaca 68700  
 taagtattta aaatacacat gggttggatc acacaggaag tggattaaat aaaaatgac 68760  
 cgaagtgggc ttcttagttt aattatgtgt catcaaaac ttaaccacta ttgccacaca 68820  
 gatgacatgt cagcttgaga tgataatata ttctgtggga ggagtggag gctgcacaat 68880

ggcagatata ttgcatgccg ctgactactg gtaagttacc tccaaaataa ttgacctata 68940  
 atattgcaaa ggccattaag ttataagatt caagtgtttc aagcatatgc agtaacctaa 69000  
 aacgggttta tatataacga atataaatt tctcatcgcc gagtaacaga cgtacaagaa 69060  
 tgagacagag agggaaatga agaatacattg tacttaaat ctttgattac atcgatgaac 69120  
 aagcgaattt caggttttag acgggatgct ttaattttgt ttattattaa tattagtcca 69180  
 tacacgggac actttaattg atgaccacta atcgtggaaa ctgcagcttc tgtattagcg 69240  
 ttttaggtag gattttaata ttatgcacag caccaactaa gcttttagac gtatattatt 69300  
 aaattttata cttaaaagca taagccatta actgacagct aggttatcat aaaaccagga 69360  
 gagcttatct ttcaacttaa cactttgatt gcatggtatt tctcttcgat ctaatttaag 69420  
 ttcttctttt tcttcagtga ctattaataa aattgtatta aataattata actcatactc 69480  
 tcgatcaggt caacttagcg aaataagat tggatgatag taatatatat gttagtagtg 69540  
 tacatcgatc tactctaaga ctaagagcgg tgtaataata taactaacca gttgataaga 69600  
 ttcatatttt acaagagggt aatttaaatt cgaatgtgta aatgtaaaaa cttcgttggg 69660  
 gagtaattta aaaatactta tgtaagtact ttgagtttta agatctgtta tgatccatgt 69720  
 aaactttcat tactcaactt acttcttttt tttagtga aaacacatta tattaatagc 69780  
 atattttcaa ggaataatt tatcatgaat ccacttaact tctctttcat ctcatctcta 69840  
 ctttctttt cttaactctca gtttttttta ttggattacc tattatatct atcattttct 69900  
 ccctttattt tttttttctt cttattctct tttctttcct ccctaacaca cccattttta 69960  
 ggatgtaagt ttatgattac tttattttca gatacatctt acaataaaga aaaaaaagg 70020  
 gataaaaaaa accatatatt actagaaaa acataaataa agatggaact tagagaggga 70080  
 tcctcatccg tcttttagtag agacaaaaac agtgacaaaa tgcaatatac agaataagaa 70140  
 caaaaatatt ctaccttaga gacaacataa ggatataatc gactctataa tattccatcg 70200  
 tctctatatg ttttaataa tttggacaca gatttttagtg acaacattag ttagtttggg 70260  
 ataattgtca acatagccca tctctaaact cgtttctatg tatctatttt aatattttta 70320  
 aaattataat tcattagaga ctgaaatgat tccgtctcta agatttttta tattaccatt 70380  
 ttatctattt caactagagt taaagtatat gctgtattat agtttgatta aaataagta 70440  
 ctcaatatta attataatat tatataatta attgggaact gtacacatgc cttaatataa 70500  
 tattatttta ttatattata tatgagtatt tatatttttt taatgtgatg gaaaaagaa 70560  
 gataagaaga aaatcaggt gagatgatac aagaatatat atatatata attccaaatg 70620

aaataccaca aacattatat tagatacatg ttgacatgaa acaagtcact ccaaaaaata 70680  
 tatatatat caatttaagc ttattgtgtt aaataataaa aacttgagca gaagaattat 70740  
 agagatgcaa gttgctacat atgggtcgata accatctcaa agaaatcctt ttggtcagtc 70800  
 tgttggaac ccaacatatac attgggtgatt tgggtgctctg tctagtaatt tgagattgtc 70860  
 tatccgctct accactttct aaacaaaaat aaaaaataaa aactacctca acaaaatcaa 70920  
 tgcattaagc ggtagattct gcatacagta gtatagcact accaagaaaa aacttcatat 70980  
 agcaacgttg ctctcgtaat atcaccgat attatttata ataagaagag aaatagaag 71040  
 aaaaaaatt ggtgatgtaa taatataaaa agaacaataa aatgaataat gaaaaaata 71100  
 taattactt tttctcgcat agaattggaa gcaaggacct ccagaaaaat ttgaaacgt 71160  
 taaatatgag tggacatggg aagatgtgat aaagatcgat ccgaatagcg tattaggagg 71220  
 ataattttta taacttttta atttataatt gttacgaggg ctaggtagtg actttttctt 71280  
 ctttttttct aattcagggg acttgtgctg ttgtgcagtg tgctgcaaaa actggagttt 71340  
 ttgtgtgtca aaggaaaaga atcagaccag aatgaatttg tgcctagcta ccctgtattg 71400  
 acgaagctac tgtatgcgtt ggggaatagc gcaataataa tgattttttt gcaacttcga 71460  
 gtcatctccag tcgcagtaac gaaaagaaga atatatagga aggttcataa tatatatata 71520  
 tatataatca tatcgtataa ctttcataac tattatatta ttttttaact ttaattataa 71580  
 tatttaatat tttcttaaaa tttatcatta aattcaaaat tcattttata taattatat 71640  
 atacaaaatc taatatatat ctaaaattat ttattatttt gtttatatag attaaaaattg 71700  
 atctatatga atatttttaa atataaatta ttgattattt ttttattgtt gttcaatttt 71760  
 aataaaattt gattcaaatg atagtaataa tatatagata taatttaacg attagattaa 71820  
 taaaaaaat atattatata tatatatata attataaaaa ttatatttgt ataatttaatt 71880  
 ccgactttgt ggctatatat gtgtgtgttt ttgttcattg tgtattttctt tctttctgca 71940  
 aaccatatag cagtataagt gttttgtaat gtacatcggg cacaacttgt gcttcttatt 72000  
 ctttttaata aatttgatcc attagctgtt ttttaaaaaa aaaattgttg aagataagta 72060  
 cgtatcaaat tttcttatat aaattatgta tcaaacgaca taaaaaatat gaagcagtaa 72120  
 aattacaat actttatggc aaatgataaa cgatatctta aaagtattat taaaaaactt 72180  
 aaagtaaaaa tatttttatt agaaggaaaa aattatatata ttacaatat tttttttatc 72240  
 attcttttag aatttcaaaa ataaatttta atttcttaaa cagatcactg gtatgcacaaa 72300  
 tccatactat atatatatat atatatatat atatatatat atatatatat 72360

atatatatat atatatatat atatgataag aacctctctc cttactactc cctatgcatt 72420  
 agttgttgac taacgccact taccaccaat ctcaagtgtc tctgactaag tccaattaat 72480  
 aaaggcacaag agcgccatcg tgccccgcac aatttatgtt aaaaatgaat acaaatcgaa 72540  
 aattagggtg aactgatgta atttttagtt tgtggctgtc caagactgc gagttatata 72600  
 taaggatgaa gaattgtcaa tgcaacttcc caagtgcaga aaattccatg caatcagtct 72660  
 tcttctttga ttttattttt cttttgtctc tttataacgt ttttgacttc cgtgaagtgg 72720  
 caaaggcaac atattttgtg tattttattt tttctagtca gaattgagat actttttctt 72780  
 caacaaaaaa tgccattcta gtattattcc aatatttata ctgtagtaat tcttaccatt 72840  
 tcgtccactt aatttgaccg gctaccattt cctgttctaa ttgggtcaa ccaccgttcc 72900  
 gttttaagaa ttttgatcca aacgcgcacg tcattccaat aattttataa tcttaattta 72960  
 attgttttcc tcttctctga ctgcgaccaa agtattcaat tcttaaagga ttgataagac 73020  
 aaaaagagta ttaccacaac aataatgcac ataaatagtg cttagcacct aagtcaacac 73080  
 tgtacacctt atcacttgtt tttaaatcaa aaaatatatt tttatactt aaaaaatata 73140  
 cgtacattgc acgggtttgg atggtagtct gtatcaaca aattaagcaa agataagtct 73200  
 aaccgcttca cggggcgaga cgagaaattt ttgtattaaa gtcattgtgt tatatgtaaa 73260  
 aaagcaaaaa aagtattagt gcattagtct agttgcattt agttagcatc aatatgaaac 73320  
 ctctctctaa gctcttaaac aagctttgtg ggaaaatgat aatgctatta gcatcccttc 73380  
 taatagagct tatcattctc attcagaagc taagggaatc acgtcccatc agcaccgcc 73440  
 aatacatcaa actcatcgag aagaagaacc ccacaatttg ctacaccaa agatttaact 73500  
 tgaaggcaga gcacgctaca gagtgcaggg tatgcttgct cgaattcgag caaggggaga 73560  
 agctgaggaa gtcacaaatc caacacacgt ttcataggga ttgtttggac aagtgtgtgc 73620  
 aacagtactg ggctacgtgc ccactttgtc ggaaacaggt gttgctgat gatgttgtgt 73680  
 tcaagcaccg tcagcatcag aatcaaccag aggctgcttc taatgggaat catgacaatc 73740  
 ttctctatct tttttcagca tttcgtgggt gcaacactta atttcacag ataactctct 73800  
 ctctctctct ctctgaatta ccaatgtctc tcaacatatt tatagtagta ctatatagtt 73860  
 aaaaacaaac cacctaatta ttttttctt ttgaaaaagg gtttaatttg ttattcttcc 73920  
 cctgccagga ttctacggag gaatattaat attgaagtac acctataga tatagatatt 73980  
 attaaagcat cgatgtacc ttttttgtat ttcactcttc ttctgcttat tatttggatt 74040  
 aaggtttgac tgatctagaa attcatattc atatgtctgt gtaagtgagc attttaattt 74100

tgggggttaga aacatccggc aattgtgatg gaggaaccc aacttcctac tgacggtaga 74160  
gttggttggt caaatgaggt agtggagtct ctcagacaca tattattcat attctactca 74220  
tattctacga agaccctttc ttagtatact gataaattgg taggtgaata agttgaaaaa 74280  
agattttttt acatccgttg acgcctttag agagaacag agacaagaag aaaaatag 74340  
atatattaag tgatatgatt taataggaag aaaagagata tatagaaaaa aaataagagg 74400  
gtcaactagg tattttaaatt ttaggtgat taaaaaatca ggactcgaat aagtttatgg 74460  
tttgatgtcc agaagcaatt cttttgatt ttcataaaag ttgctttaat gttaatatgg 74520  
attaaaaaca aatattaatt aacgatatat taaaatgaaa gaaagttagt gaattttctc 74580  
aaaagtgaag aacatagaat tatacagtta cacatgcaac ataaaaaaa ttaaaacaat 74640  
ataggtttag gttgtaattt gttgaaaaag attatgtttg gatcttctat gactatgagc 74700  
acataaattg tgcttagcac caaaagtcag taaattttgc ttaaaaaagt atttcaaaaa 74760  
tagcaaaatg atacttaaatt ctatttaatt gcataaatcc aataaaaagg gttaaagaaa 74820  
ttaacatcct aagttaaata tatgcttaaa aggtaagcgt aaatggtaat gataattatt 74880  
gataagttgt ctttcttcgt tccaccagct taaattaact taggtcatgt ttgtcttcac 74940  
atagctagga aagacataag tctcatcgg tgatctcaac ttaacagact tgacttagaa 75000  
atagacgttg ccacgccttt cgatcatttg ctaatcatgt ggcagaacca aattaaggag 75060  
aatatgtagc ctaggacatt aattattcaa tatggttagg aggtagatgg acaaatccaa 75120  
tatcatagga cacaggggat gcatacaact gacacttttt gggatcaatg aatcctcaat 75180  
tggtttttcg atttctgtag cagtaaagtg cgataagaat taatttgtca atcttcatat 75240  
tttcataaat aaatttgatt atttattctt atattatgat ttgtttgtga atgaacaat 75300  
ataagaatgc atcaaacagt gttgaaagtc gttaggttta atgggtgact cctatcatac 75360  
tgcaattagc taggttctta actagtctct gacatttttc aacacatttg ctgggacagt 75420  
tgggcaataa accaatgggt tttgatgtaa aaaaccaatg acaacaactg gaagtagggt 75480  
tactcttcac atgcatgtgg togatgttgt attattgatt atatcaaat ctattacaat 75540  
tttctataag ttttcctttg cgtattggca gaagaataat caagaagggt tctcagttat 75600  
gaactttgca agtacaatag atgcgtgttc cctaactcct gattacagac atttccccctc 75660  
aaattgttca tccaaaatga attatttgct ccttcttttt ttgagttgaa agagacataa 75720  
gttcaacgtt gactttattc ctctacatgt attcacatca catacgtgat aaatgaattg 75780  
tagtacttac acaatttttt tataatatta tttttatttt ttcttattac atacattaac 75840

tatatattca cagctcccaa tctcgctctt ttctttctcg gtctctttat tatacattat 75900

ctgttatgaa aaattgtata aacataattt gtcgtaataa atttcaaaat acaaatgtga 75960

actgttatat atttagatga ataattataa ttctcatctt ttttttttaa ataatttttc 76020

ttataattaa ggagctaaaa atcactctta gaatgaataa tgtgagatgc ctaccaaac 76080

tcaaaagcta gttttgatac gatgttaaa gtcagcaaa acatctaaaa gattttttta 76140

aaaatcaatt aataaaggga tgacctatcc aattatataa acacttatca tattcattaa 76200

ccaccgata cggctctcaa cattcatctt agaaataaaa gcttcttgta tatcaccaaa 76260

agtttaacaa aattttgtgg acaatgaagg atcattggat aatcactgta agcaacttgg 76320

tcgtagtaga gtcagtacta gttttttatt atctactcga ttatactagc aattttctatt 76380

tcattccatt gcattgtatta aaatactact agatttatatg aaatgaaag tcaactgggt 76440

ttattatcat agtgggagtg gatttagtgt tgaaatggta gctcaaccga tcgtagtgac 76500

ggttaagaag tgcgggtcaa gtagtgaaca caataattga ttacaaaag catatgtact 76560

ccaccacta actcttgcgt gctccttggt tcaacaatt aaattaatat aagggatttt 76620

gttagtgttt aggatatgtg ttaagaatc aatatatata tatatatata tatattgtat 76680

aaattataag agaataataa aaaattatga atagtataat tttttgtttt tcaattttta 76740

aaattttaat taaaggaaaa ttggttatat tttaaattat tctttatgta gaagaaaaat 76800

tgataaatga ttttgatcca gtatatgtcc caggggcaag gggggaata acagaggcta 76860

tagctcttca agtactgatt ttcaggctact ttgttgagtg gtgattctaa aatcgtgggt 76920

cagtttgta accatgcgtc aacacagtac tcttcatctt tattagtcta tcgttacctc 76980

tattcgtcat tgggtggaca agaactgtga ggttcatatt gtgcaagttt ttgaaaaaac 77040

aacctgaac catggatgct gattttttta gcatcagaag ctcaatttta ggagatatag 77100

gagttcatat ttatttctaa accatcagaa cttgattgac gcaattatga tgagacttcc 77160

ctaggcctgg tctgttcaat ttgagatact ttctttctca attttttttag aagaaatagt 77220

agaaaaataa ataaattatg agtctctgga agaagcogta ggaatataaa cttgtagatg 77280

catccacact ccaatctcaa taataacaaa tattccaagt gaagtaatta ttgaaattat 77340

gtctcatctt ttgggtgaagt cgtctctoca attgagatat gaaattttca tgactgtgag 77400

aagcgtagcc aaggtgccag gttatgtgtg actacgatgt cactgacatg aactatctca 77460

aacaatgcaa aaattgtccc cttaaaggca ataccctctt tccccctt aaaatttggc 77520

tttcaaaag tacttttttc aaaatacatt aatttttaac atggctagaa agtatttttg 77580

aatattataa attttaacac ttccagaaaa tattttctga aatgtgtga aattttttaa 77640  
ctgtgtttga aaagtagttt ccaaaaaaga gatacttgta gaagaagaag aatatatata 77700  
tatagtatta gtaaatctgg gaggtgtaaa taacagttgt taccacctaa tcttaggctg 77760  
taacttggtc cagatggat tggattttaa tctcaaaggc ctagtacagt cgacagatat 77820  
atggcaagaa tttttagggt agcacactta agactatgtt gaatatatgg aataaaattt 77880  
tccaacttca aatttcaaac aataacaaat aatccatgaa acaaacgact cctatactct 77940  
tattgagtgt tgatcattga ttttaagtact tacaattaat aatacatatg aaaagaatat 78000  
atatactatc gaagaagtga ccagccataa aagtgttggg ttccataaag cctctctttt 78060  
aggcccaata tcataaaga acgggtcgca cagggtccaa taaaacaata gtgtcccaag 78120  
agggtggcca gcagtagaca aagagacac aaggctttgt tattctattc tcccatcaa 78180  
gttcacacc catattttct tttatttttt taacaaaaat atccttgaca gacacatttt 78240  
tattatgtat tgtacaaaa aatataatat aaaaaaaca agtatttatt gtgttaaatt 78300  
ttcattaaaa aaatttaaca tgactgaac ttattataac ttttttcaaa aaagtataa 78360  
tgacactatt attgtgagtg aaaaaagggt gtgtgttgta attgtgagaa agtaagtga 78420  
gaatgtggcg tgagaagaaa gtgatgtagg aaatggact tgggtgagta gaaagtatgt 78480  
gttactgtaa atgtaattgt gagaaagtga gtgtggtgtg agaagatatt gatgtgagaa 78540  
gatgggtgtg gtgagtgaga aagagtgtgt ggtgctgtaa taattgtgag aatgtgagtg 78600  
tggtgtttga agagagcgac ctgggttgtt gagaacataa actttttttt gttgttttag 78660  
tttaaaataa tagcaataca acggaacat ttatgttttt gtgggttaa atatttttta 78720  
aaaaaaata ataaacaaa aactagtctt attttgttca atcaaaatat atttttgttg 78780  
tgtattataa aggggtatca tacattataa tagaaatata ttttcattat aatagaaata 78840  
tattttgatt ataataaatt cacaacacaa cctagttttt aaataataa taaaaaaag 78900  
ctgtggtaag aggtggaggc aatgggacct ataggggcca tggcccccct cctacagat 78960  
ttttaaacct tctaatatt atgtataatt aattaactt acttttgatg atattatttt 79020  
tcctttgggt ttggtaaga tgtgaaatgg tgggtcatga gtgtaggatg tgcaacacca 79080  
accaggtttt acaattataa gtctttaggc atttgatgtt ttcagactct gtttaattgg 79140  
tctacactac ttgagttgtc aagttgtaat cattgccatc ttaataaaat tcaactttgtc 79200  
ttaaaaaaaa ttaatggctc tgcaaatgtt aattatttc acctcacata tgtttctcatt 79260  
cctttttctt tctaattgtg gtaaagggtt ttgcattgca ttccattaac ttgtagaaaa 79320



actagttttg aacacataaa acaactagtt tcttttaata ataagtctta ttagatttat 79380  
 agtttttgtt ttactaaca atacatttat atttagctaa catgtctaac tgtttcttgc 79440  
 caatttttat aatacctata ttcattgtaat gaaactcctt ttctactaaa ttaattttaaa 79500  
 tttgtgtggg ttaatttttt cttttttttg gttgagagat gtttaagaag aaactataac 79560  
 atatagtaat ttataagag aaagaaaga gatgatgaaa aaacacatat ggtgtctaat 79620  
 gcatttactt caaatttaga tactatgttt ttattgacat tattggattg aaacattata 79680  
 ctttgataat aatgagatga gtgacttatt gttgattgta tttttgaagt ttcactctta 79740  
 ttttgaagat aagagatgct atctttttgc acactttttt taggatattg ttgtcagttc 79800  
 ttattctctt tctatttttt ttgttattca tcattgtgtg catgtgtatt gttggcactt 79860  
 attggttaat cttctttact caatataatt atatatctct tctgcttata agaaaatatg 79920  
 ttaaatttct tgcataatatt atgcattgta aaatttaatt gttttttttt tctaaatcaa 79980  
 gattttttat tagccctcta ataaaatcct ggctccacc cggctgtggt tgtaattgtg 80040  
 taaacttaga gaaaatgtgt tgagaataac acggttcagg ggacagtaat gcaaaacatg 80100  
 tacgccact ctgctaaaaa agggcactact aaacacaggtc tgcatatata ttacttctgt 80160  
 ttttcatgtc ctgctaattt cttgtttaat caaaatatct atactaattt cctatagagt 80220  
 ttacactaaa atttatttat atcaatacat atccctccaa ctccaagagt acatgtatgc 80280  
 atgtacgtgt gagtaatgag ttctttata acctctaga atcccagaca atctaataca 80340  
 tagtttaaaa attacaaaac ataagatcag atatgtcaaa gtatatatat aaaaaggggc 80400  
 accggagatg ttaaacgggg aacacatttg cctttaattc aaacataac tatgcacagt 80460  
 aacacagaga aacatagatt acaaaaggaa aaaataataa aacaacaagg tgccaacaac 80520  
 agccaactat tgcaaccagc aagctgaagt tgttcaagca gaaaagaaac tgttagaagc 80580  
 agcaattgca aattgggttt tgggttccaa aaccactgtg acaagctcct tcttctgcac 80640  
 ctcaatcttc tcttcagaa ccaacaaaga agcacaagaa accacgttca agttcctctc 80700  
 ttgtacagcg ccagcctctt taaaattcat tgtcctcctc tcattattct ctgctaccaa 80760  
 catagcaata tcagacatgg ttgatggtag attctctcac tcactatata tgaacagatg 80820  
 ttgttttggg tgggttaagg ttgctcttac taaaaccttg ttcaaagcaa aagctattta 80880  
 tagcgacgga ttctacgtct ctacgcgtgg cgggtcaaca gtttcggata aggctcaagc 80940  
 cgtgttaggt ttgcattatt aaattattac tcattactat actttttttg ttgggggtgg 81000  
 tcataatcat taatgtgtcc ttctgttaaa aattctgaag actttccgac gaaatataga 81060

ggactaacta atgtttcttc aacctaccca tgtgtttcac atgccctgaa tatgatagga 81120  
 caatttggac agatctaagtg tgggagaatt ataaaaaggg gactgaaaaa aagaaaatgt 81180  
 tttcatatata aatgtttgta caaaaaattg tccaataagg aaaataagat gatcgatgca 81240  
 ataaaaagaa aataataata ttataaaatt attgtttcaa tcaatttgta aataatagaa 81300  
 ctaaaaagga gtttaacttt taaagccaaa gatggtaaat gtaattcact ttttaaatga 81360  
 acaattataa ttcctttttc aaaaggtata tttttacttt ttttcatcat taattatatt 81420  
 tttacttttaa agaagttaa ttttaaaaag tgaaatgttg tggtaataat ccaaacaaga 81480  
 gttttgcatg atttgtaaaa tattccattc tctgtgatcg gtcatacaat tggaaaaatga 81540  
 caaaatattg aaacgtggga cacgtgtaac ggtttctcac tgtatcgtga gtatagagac 81600  
 aataagcaat aatgtccgac atgcacatgc ccaagcctgt ccttatcctt tttttttttc 81660  
 agaagggatc attaaccttg gaatcctaatt tcctaaaaat gagatacaga tatacgtttt 81720  
 taaactacg gattggtccg atgggttagtc atattgaatt caccttatgg agaagatata 81780  
 tttagcacia aattttgtat ttatgtgtga aaactctcca gataagtga agatacaata 81840  
 tgataagtgg aattagtttt ggggttagta gactgaagga tacccttgac aatatcaaaa 81900  
 aacatatttt aaagtagaaa aaaatgaaaa gaatttaaaa aataataaaa gtaatgtatt 81960  
 ttttggtata gactacatgt ttaggattta atatgcatat tcagaaatta aacttgagat 82020  
 atctaattaa attgaaataa tatcacta gttgatctaa gcgctcataa taataaacgt 82080  
 gatgtaactt cttcttactt tgatgtttta ttcaaagcta aaaataattt taataattgt 82140  
 cttcatatca tcaataaaat ttctccatgt gtttcagcta atgataatga ctccctctta 82200  
 gtgtctttct ctattgatga atttagatct acactctttg aaatggatcc agataaatac 82260  
 acatgaccaa ctgggttgaa accaaccttt ttaaagaaat ttgggcatct ttgtggatct 82320  
 gagctcttta attctagagt ctcatggctt gaaactagcc aattccacc tcaggtaaat 82380  
 aaaaacaaa tttttctcat cctaaaaaa tctaatacca ccacatgaa agatttttga 82440  
 ctcatagtcc ttgcaatgt tatatacaaa attatgtcca aagtccatgc caatggcctt 82500  
 aagcttatcc ttccaaaatg tatctctcct aaacattctg cttttatcca agatcggtcc 82560  
 attttggata atgtccttat agtctctgaa attatccacc acataagatg caaacaacaaa 82620  
 ggaaactat gagaaattgc tttcaaaatt gatatcagca aaaaaattta ttgagttgac 82680  
 tcgaactatc tttttggcat catgggtaag atgggattta acaaaaaatg gataaactga 82740  
 atgagaattt gtttcccatc aatccaattc taggtggttg ttaatgggga taatgtgggg 82800

gaaatttttc gagggagaag acgaaagata gaatgactca ctaccatcct acctattcat 82860  
 tatttgcact ggggtctttc aactattctt aaaacacggg gaggtgagag gtgacataca 82920  
 tgggtgtgaag gtgtgcaaa ggaactcccc cccccctctc acttattatt gctaagtatt 82980  
 gttttctttt ttgtcgccct aaaataaatg aaattttcac cctttaggag atcctagaca 83040  
 cttatggaaa agtatcaaga caacttatca attttcaaaa gtctgaaatc tttttgagtt 83100  
 cgattattca acaacaacaa agacacaaaa tctcttcta ggtgtcacta aaagcatttg 83160  
 atccaacaaa tatttaggac ttccttcaat cattggcaag agggaaaagc tatttttggg 83220  
 ttcatcaaa atcgtctttg gaattggctc aatcattgga ctagtaacaa actctcaaaa 83280  
 gcgggtaaa aaatcctatc aaatatgctg ctcaatctat acccttttgc tgcagtgtg 83340  
 cttttttgct ccattccact ttaaaagatg attttgaaga aaatgatgaa gtgttttttg 83400  
 tggggcaata acaacaagg aaatgaatca attggctcag ttggataaaa ttattataaa 83460  
 aaaagaattt ggaggaatga aattaagaca ctgtacgcc tttaatctta ccattgttagg 83520  
 aaagcaaggt tgaattatca ctaccaacca agatactata atgttcaaaa tcttcaacgc 83580  
 taaatatttt tcatcaatgg atttcttggg tgtccacctt gggcacaaat ccatttttgt 83640  
 atgatttagt attcacgctt cacaggtaat ggtaaagcaa ggtgttattg gaggattggc 83700  
 aatggtgatt ctactcatgt ctggactcaa ccttggctct aacaataatc ttaatccctt 83760  
 tgtctcaacc accccaataa tggatatctc cctcttaaa tttctgacct tattgatcat 83820  
 gctaattgcc tttgaaggaa tgatattcta gattcaaaat tcaaccaaga ggatgctcaa 83880  
 aaaattcgtt cctaactact attgaatct acggagagag atcaactaag ccggaaatat 83940  
 tcaacgaatg gtgagtacac tgttaagagc gcatatcatg ccattatgga tcatttgcta 84000  
 gacaatagcc acctcaaggt taaagggaaat tggaagacca tttggaatct tgttgtgctc 84060  
 caaagaataa agcacttcac ttgaaggctt cttcgaaaat gcattccaac aagatccaac 84120  
 ctgcgcaaga aaaatatcct ctatccctta catttgtgtt gctgtgatc aaatattgaa 84180  
 aatgagtggc atctcttctc tgccttgatg catgccagac tcatttggca agctgctgac 84240  
 ttatggaatc tcgctgagca tagatgggac aatgcagata atttccatga ttttatgttt 84300  
 gattttctat aggaaagcca tcttttatgg cctctctctc tgctatgcag catttggcag 84360  
 aatggaagca tgcccggtgc ttgcacaacc aataacgtcc acctacaact cataatggca 84420  
 cagcatcata aattctctct caccaaagac atatcaaatg taatttggat gctgccttac 84480  
 tcaaggatc tcaaaccttt ggcataggaa tgtgcttgcg ggattctcag ggaattttta 84540

tcaaagcaaa atctgctaca attcaagaac atccaaaacc tgaagttgca gaggcattggg 84600  
 ctctacacca gtccatcaat tggataaaaa gatctacaac tccagaacgt gattttttgag 84660  
 ttggattaca agacggtggt tgataatttg atcagttggt ctacaggaaa atcagactta 84720  
 tttaacaaga tgtagagcct ccattctcttc tatcccaaac tcaaaactcaa ggggtgagtgt 84780  
 tataaagaga cgaggaaatc aagcaactca taacttagca agagtttctc gattttatgt 84840  
 tagctttggt gtcttcaatt acatcccaac ttgtattggt tctcatgtta tgaatgaaat 84900  
 gtaataacct atgaagtacc gataccaaca cggacaccgg acacggacac gtgaacacct 84960  
 gtaattgttaa aaatatagaa cgtagtattg gtgtttggaca ctggcacgga cgcgtgtcga 85020  
 acactagaca cgacaagagg ttagaataat catgcttcac agataataac tatgtttttc 85080  
 atttaaaaaa aacatgatat ataaaatatg ataaaaata aatataggtt tgatagtgtc 85140  
 ttcagttaca tcccaacttg tattattttct catattatga atgaaatgca ataactatat 85200  
 ttttcttttt aaaaaaacat gatataataa tgtgataaaa aaataaatat aagtttgata 85260  
 tatttttatt ttgtcaaaact actttcttag tgtacaggat aaaacaacag aaaaatacaa 85320  
 gaaagtaaaa taatagttag aaataataaa ataattggaag atataataa gatcatacaa 85380  
 acatccggtc aaaacaaaa gatgaacaaa aagagagaaa aaaaatgtaa taaaataaaa 85440  
 gctaagaataa aaaaaaagtt acatgtatga atatcaacat cctttggcct taatacgtgt 85500  
 cacaattttt caatttgatt cttacagtaa aaaaattact aaattttatt tgccaaaaat 85560  
 tatcaattta attcttatca tcatatgtca atcaataacc attgatctcc ttacctattc 85620  
 ttaatttagt ttttataata tagagcattt ctaattgtat cgctatatat atatatatat 85680  
 ataaataatt gtatcgctat aattttattt taatttggta ccgtaataa gaaaccaaat 85740  
 ttagaataca ctattcccat ataaaatcat aaaaaacaat tcacacaagt gataaaaatt 85800  
 tatgtctctt aattaaataa ttcaagattt aaatcctaatt tcacctttta agtatgaaat 85860  
 aaagataaaa atatacaaat tttagtagt ctactagtcg tacaaacact aaattatata 85920  
 tttaaacgca aaaaatgtac accttatcgt tatattatag ccttgattta taacaataat 85980  
 aattacgtgt catttgtaagt cgggtgatga tgaatgata ccatatatgc aacattatata 86040  
 acgtggctat tttcgtaaaa ctccaatcca aataaaaatc cgtgttctga atcagtttct 86100  
 gactcgatca caaatccgtt atttttttag gaaccacaac tccgatttat ccataacaa 86160  
 taagttaata ataagtaaaa tggacatgag ttttagcaaa aaaaaataga aaactaagc 86220  
 atactaaaac attatcattt taagaaagaa agaaaaaga ctatcataac tgattaggat 86280

cataaaattt cctgaactc caaatttctt ctttgtcaa aattatcaat ataatcctct 86340  
 ttgttaaaat caacagcgtt atgggatgct taattatcaa ctcaagactc agttaatgac 86400  
 gttaaagcca agaagattca cttccattat gggggacaac aaatttaatt taacagaagc 86460  
 aacatcatca taactaaaa aaaaaaaac taaactgatg agaagagaa tgaatacaaa 86520  
 actacgacct agaacaaca caggaccaa catatatttt tctcactctt aaacaatatg 86580  
 cttttaaaac ggggcaactg tacgaatgaa attctaccgc tagaattctt tgattagtaa 86640  
 agtaatccaa atcaaaacca aaatttaaaa caaatgggac aagcaatgaa taaaacctga 86700  
 ttacccaagc agaattctat caaggaaaag aaatgatgta cctgtaacct gttactggtc 86760  
 attaaaagaa tcctcttcca ttttctttac ataagtagac cgtgtcagac attcacaaat 86820  
 ttccacatat cagcgggtta gcacaagcgt acaactctag ccacaatgaa ttgaatggct 86880  
 attacacaat tatcaattaa ctgcggtgtt tcatgaaaag gaccagctta ccaaaacatg 86940  
 gggagggcca atgtatatct acttccattc aagtcataa gtagatgact gatgactctt 87000  
 tcaaaactcg actcatagta ggcctaattg tttcaactgt tgaatatgct ggggtggaag 87060  
 agccggggga gaccaatccg atcctgtcaa ctctccatca acatcttcaa ccacatactc 87120  
 ctatgatgag aagcaataat agtttcacaa tcaacaattc aagacgacgt cagaattttt 87180  
 cagggcatac tgatgttttc ataaaacatg cagttaacac atggggattat cagtactgat 87240  
 aactaatctc aaactgcttg tgttccatgt caaatgtaat ccagaaagag gccgcccaca 87300  
 agccagcctt aaggcatcaa aataccagac tcaagtctaa gctcaattac tatcacaccc 87360  
 ggataagcga acgagagttg caaaatcaaa tctaaaacac ttgtcatatc agacagacaa 87420  
 aaagactaag gagagagaat acatccaaca aaactcttca aataagttga tgcagacatt 87480  
 gttagcaatg aataaaaaac tcagatggta aacttcaact acaaaattgaa tgcaactcca 87540  
 aacggtgaag acccatcaaa tccaaatctc gaaactgtgg gtcaaatcac atcctctcca 87600  
 gccaatcttg tctgcagcca aatcataacc gttgattggg tattgcacat taattcatgc 87660  
 agagagaaaa tattgataaa tgggtccagat ttggctgca cagaacatcc tgcgtgcagag 87720  
 gcaagagagg atcctgatcc acagtgacat atctgacaac cgtatggcag cctttctgtt 87780  
 acatcagtgc ttctagcac cacagcccca cctagctaa attatacagg caatggaaga 87840  
 gcaggagggg attgcataat gcatactcat gtacttggga tgggtgtgtg agtgaggata 87900  
 aacttgagga attgatttgt aacacagctg cacatgcagc aggtgggacg aaggagtttt 87960  
 gctttgagaa cagaatcagg ggaaggggga gggaggagga agtgaggcaa gcatgtactg 88020

agcaagagac agagaagaag aaaagtagac ccatcaaaga gtctgacata tttctgtatt 88080  
 tactttctttt ggctgagctt ttatacaatt aaagatcctc atctatataa tggatcata 88140  
 ttcataatcag tagcaaaagat gcatattatt acgagagttt ggaagtactt tccacattat 88200  
 aattccagtt cctctgcaaa tgtcaataat ttattatact agaaaatcct ttggaagtca 88260  
 atcaaaaacag ctacatcaaa agcaatgcac gttatataat ggtaagctag tagctgaaat 88320  
 acattttactt tcgcaaaaac atcaggcacc ataaaaatcc tcagtcaata tggctctaata 88380  
 gcaagagaat ctaaaaaaca agtcaattac aaaagcagaa aaaattatgg ttgagattct 88440  
 tggctattcc ttccctggta aatgtttctg tgaagatatt atataggact taaatgctaa 88500  
 aaaaaaccca agtgatggct aaatgggcta attgcataata tcagtttagga atttagcatg 88560  
 acttgacagc tgaattgcaa caagaattga cagctcatag caatgcacca aaagcacaaa 88620  
 agatgagaaa gtaaaccaac cacaccttgg ttaacatcct ttttgcaagt atgtgataat 88680  
 gacgttgcca tatccgttga cctaccatgg tagcaaccaa tacactgtag aatatcccaa 88740  
 ttacagtaaa aagtccaga acaatcaaag ccataatgaa taataatggt agccctgctt 88800  
 cccagcacc ccccaaaaac ccaccacatt cagttgccat ggttccacaa ctctcaaagc 88860  
 atgtggtgca atcagttccac atacaagaag tgccctggcaa atgacagtct gcacaaactc 88920  
 tgcaacaaca tacaaaaaca aaatcaaaact tagtcacgaa tccactgccat tctgagtcac 88980  
 aaacttgctg ctaataatac ctaattcatt aataacacga cacagccttc tttctctgtt 89040  
 tagcataagc aagcatgaat aatttaacaa cggaatatta tgtcttccaa aaaaagttta 89100  
 ttaaaagaag ggtctgtgat atagtgtaaa ttaaatcttt atctttatat attacactcc 89160  
 aagcatagtt tctggcatcc tacaattaag caagaaaaca agcattataa cgatgcctca 89220  
 tgcagatgct cggagtttca tcatgtgttc tgacaatatc agaggcttgt gaatttgtaa 89280  
 aatcaagggtt attaccaggt ttgacagcaa caaagacata gttctctaca aggcgtgagct 89340  
 aaatcattgc gtactcttcg atcatagcag gtaatgaagc acccagaaaag gccaaagcaag 89400  
 gcgaaaaaca atagagctcc tgcattaaaa aataccaaca aaacatcaac acaaaaaacca 89460  
 gtatctttttt aaatgtgaa caaggaaact atgaactcaa ggtgatggag gaaaaaaga 89520  
 atttgacaga aataatgtgt tataagcaat ttgatgacac cgtttaagct aaacgcaatt 89580  
 taccacatat atagtaaaag ctcatctcac tatcaaaacc ccagagaaga cgaagccaat 89640  
 attgctggta accatctatt agataaacca aatatgccag tgaagcaatg acctgcatat 89700  
 aaaggtaatt aaggggataa attggatcac ctatatgcac tcataaagta aaaatggata 89760

agtagatgga aagattaatg gaaacatatc acctacaagc tggacagaca gaaaaataaa 89820  
 caaaatatct ctggtgacaa aaaaccgaaa ttctagggtg cgccatttcc tatcagcggc 89880  
 aacatgaaca cgcaaatgat aaggagcctt gcacgtagtg cagtgaagcaa aggcaaaccc 89940  
 ttctctgcaat aaaaaatgaa gtattatcca tatatttctt aggtcaagca ccgcaagta 90000  
 catcatttta cacattgtta tttaatatgg aatatccagg tgtttctcca tcaatgggta 90060  
 ttactcagt accttacacc ccacaatgca ccatatacat tagaggattt tttttccaaa 90120  
 gaaaaattag cttgttgtct acagttcttt tcatcactag actaagtcac aacaattcat 90180  
 tttaagata attaattgtt acagtatcat actatcatta aaaaaaagg aagaatagaa 90240  
 tggagacta aagatagtta ctaacaaaa atgtgtcaat ggcaagagat acaagagcaa 90300  
 ggaaacatg ttaaaagcca tggagagtaa atcaaaactg caaacacaaa ttgcctaaaa 90360  
 taaaattgac aaacatgtaa actcttttca tcaccgaata gggtaatggg tagccaagag 90420  
 agagagataa ctgaagggtg agaatcagca gaaatcatca gaaagatat gaaatgttac 90480  
 caagtcaagt acaggctacc aggggtggat caagagatgt atgaaagggg ggccaaattt 90540  
 tttttatata attattttaa tatttaactt aataaatgat tatttaataa attatgaatt 90600  
 tcacaaaatt atttattatc aatttcatgc gtaaaaccag agatatgatt caatacaag 90660  
 gaaattaagt aaacactaac tcaacattct aaattttctt tacttttctc tattaatatt 90720  
 tttttaaaaa gaaaattctt ctcacatata caaagagtct tgaggagggg agggagggca 90780  
 atgccccccc ttgaatcgtt taccacaggg taatagacac aggttagaaa aaatatcatg 90840  
 agatcaacaa atgaaaacat aacatagatc atttgattaa ttgattaacc aaacaaaagg 90900  
 ttaccgacc accaaaaaat agccttgggt cagttttgca taaaagatag agcttgccat 90960  
 cttactatc aatcacgaat catcatcact tataaagttg tgttgttatg taataaatca 91020  
 ctaatcatta ctttaattgc tcgccaatga tccaacattc ctgatgcac atattttgat 91080  
 gtacctttgc acttacaagg ggctatgaaa tctctacctg ggtaaaaaaa aacaagggaag 91140  
 ccaactgaaa aaaaatatga tgacacatcc aagcgttgag gataaacatt caacaacgaa 91200  
 gcttacatag ctgcaatacc caacacgtgg aacaataaca gtcccaatag caactatata 91260  
 cataattgca aaagaaaaac caattcaaat tccaatcaaa tgatccagac gttatogaat 91320  
 tattcgaata caataattca ttacttttgc aaccaaaaat cagctgagct tgaccaataa 91380  
 ccgtgcttct taacttaacg cgattcaagt tgttataaaa tgaattgaat aaaccgaaca 91440  
 ggtaaaaatc gaaatcacga cttattttat ttagtatata gaattaagaa tagattgttc 91500

tcaccacatc gaagtgacgc ttgccctaac ctgaaaagtg aaactgaaaa atttaaaaag	91560
aaaaagagga tgaagtga attaccgtca gtttcgagac aaattcggca ctggatttgt	91620
tcggaggagc cggcttcgag gtcgatctcg gaagggtcag cgaggggaag gggaggaaca	91680
agaggagaag aatctgagtg atcggccatt gaaattcgga ggtacaagt gtcaaaatga	91740
attctaattc ttcacatca tcgttagctt agattaaatt tagaatagag gagaagaaga	91800
atctttcact gatgatcctt ctctttctct ctctagaact gccaatgctg aattgatttt	91860
tcaagttttt gttatcacag acacaaaagt tgcgagacta cagtgaacc ggtgaattcc	91920
caaaactaac cgcgaataac aatcaaatct gcgttggata ataaaaatca atgaaaaagg	91980
ttataaaaa ttaattttt aacaataatt gtaatcattc tgatttaatt acattcaaaa	92040
caacttttta tcaatttcaa aagtaatta tataagaatt gaattttttt ttttaaaaaa	92100
aaaattagcg tatattttt agttatatat ctaaacattc taacaaatta acctccgtt	92160
aattttttta cattttaaca attgttgata aacaaaattt aatttaaaag tattactagc	92220
ttaatgttag cttttcactc ttaattataa ttttatgaaa tttttcatta attttttttt	92280
ttacttttct gtgcattact acattctgtc ttttattttt aaaatttcac gaattttatt	92340
catcggtttt tgactattta ttgcattag ttgactttat taactgctga ggtaaatctc	92400
attgaacgac actgatatgc tggcgcccaa ttaagataa aaatactcga tttttttaaa	92460
gtaagaaca aaaggtatga ttaagccttt agataatagt aaacgtatag gagatagggt	92520
ttaaattctt aaaaaagact ctttcttctc actaccatta aagtaactaa atcattcggt	92580
ttgcttgatt tatagataaa aaaatttata tataatatta atttttatta ttctaattta	92640
aaaaataaaa atttatataa atattaataa aacaaaaata ataaaaataa gtgcacaatt	92700
agacatagaa gttgttatgg aagttcttta ttgaagcttc aactgaatt ctcattataa	92760
ctccaatac aaaatactca attgtcctaa taaaaatttt aaacaacttt tagattaaga	92820
gaaattgaga aaatttacta aaaaaaactg ataaaaataa ctaaaaataa gtagtatttt	92880
ttaagtaaaa aaaaaaactt aaattaacat ttggatcact aaaattatca ataactaagg	92940
tgaaacaatt aaaaatacat caaaaatatt acattgatat gtgtatggtc aaaataattt	93000
gtttttctaa tcacgatatt ggcagtgaat aattcttata aaaaataatg attaatcttt	93060
gttggaagg caataaatc atataagata agaaaaaat ttaataatatt ttatttcata	93120
cccaaatca taattttttt atttatagtg aaatatacac ccatagacaa taagtaccca	93180
aaaataaaag tagacatagg caataagtaa atttaaaaag ccagtactcc acgagttttt	93240



ttttttgttt tttttaccgg aaaaccgaac caaatatccc cccaaactaa actacaaata 93300  
 ttatcgggta gtaaaaactgt aaacgtgccca agattgtatc tttttggtag tttttgtagc 93360  
 atttacttcc aaggaatgaa gaaaagatta ttctgtagaga ttttgggctt cggccatcaa 93420  
 ttcaagtttt tgggctcatt accagtcatt ttatcatacc tttccaagga gggaacaatt 93480  
 attgttattg acactatgct aatagcaatt ttatcattct acattccctt ctttttcceta 93540  
 catctctaac ccttcttctt tctctctctc ctttccctt cttatttcat aataccctt 93600  
 acccttctcc cttctctctt cacttcttat ttcttcatta ctttttccaa ctcttgtctt 93660  
 cactctctgc tcctccacat cacctaattg tgtatttttt ttttaataca cagaactctg 93720  
 tttttgtgtt gttttttaca caatgaaac atgcatgatt tgcctttttt ttatttttta 93780  
 gaattgtaaa ttttttaatt gttctaaaa atacatattg tatagcaagt catttaatga 93840  
 ggtccttaac aagtgaataa tgagacaaaa cataacaatg gaaattattt ttgcacaaaa 93900  
 aaattataaa acagaaacat gtttctatcg tgcattgata taatagaaac atgtttccat 93960  
 tttatgaagt gcatcagata aaaatataac aacaaaaaca tattttcatt gtacattaat 94020  
 agttatacaa tggacgcatg ttttcatagc gcaattgggg gtggcaataa aaaaaacaca 94080  
 acataaacat gtttacattg tgtaactatc aatatacaat gaaagtatgt taccattatg 94140  
 catatgggga taaaaaatt taacaacgaa atcataattc cattgttaaa ttttttttaa 94200  
 tcccatttgc aaaacaaaac catgtttcta ttttggttta tgtcaggggg caattttggg 94260  
 ataccataaa aagggtactc cactggtagt gccataaca aaagtaggca gtgataatag 94320  
 caattccctt taaaagaaag catggatgtg atggaactct atttagtatt ggggctgcta 94380  
 ttctaccctg caattttgca aaattacttg cctgccctc ctccctttcc tgcacccac 94440  
 catcttctct tcttctcttg ttgcacttga atctgcgatt aggtagggtt tttttttctt 94500  
 catactttta actgaaactg gttataattt tgttgtattt ccaacacac agaactatag 94560  
 aattttttat ttataaaaa acaacaaaat tataattctg ttagcgtgtg tctgggttga 94620  
 gatagtaact taattcgaga gattattatt cagccttacc gtgatatgat gccatcactg 94680  
 catatatcat tcaacttat atcacgtctg cattatttgt ttcaattaat catgtgcoaa 94740  
 taatgctcaa tttaaaattt atttatttaa cgtaaaaaaa tggcaaaaaa taattataca 94800  
 cagaagaaaa attaatgtac tctaagataa tctattgata acattactac ttttaataat 94860  
 tatttttcaa tattaatttt atttatttaa ccattaaatt aaggtttaaa tattttttta 94920  
 gttcttatta tttaatgttt ttttttagct taaaaaatta tatttttttt atttttagta 94980

cttataatat ttttaaatat attttgaaca gtaaaaaaat attataaaa atcatatgaa 95040

tcaaaaaata aataacata attaataagg actaaaaata aaaaaaaat ataataataa 95100

aaagaaaaaa agtactaaat tacagcacta aagaaatatt aaatcttaa ttaaagttgg 95160

gttgaaaatt cattaaatta aagttaattt tttagcaata aaatatacat ttttatttaa 95220

aactaaattt atttaatact tcatttgacc aatttaatat aatataattt aaagtatggt 95280

tctcttcatt ataaaaatag aaaagtttga tttttgaccc aacgaaactt gtgtgatggt 95340

ttcaattctt ctaaaattaa aatcttataa tatgtttcaa agttaagggt gattaaaatt 95400

gaacatatat gtttaatgaa tataagttca attttacaag tatttagtcc ttaagtggag 95460

attcatggta tatgtatgtg cattttaaat aattgacctt aactctagat caaatatga 95520

tatattttaa ttttacaagg attaaaaaca taaaaaata gtatggacaa aatcaaaact 95580

cttttaattt tataaagatg aaaaatatat ttactttat aatatgtata tatatatata 95640

tataatagtt tttttaagaa aagtataata aatattaact ataagttttt ataattatta 95700

agctaagttt aatatattca agaattcaaa taattaatat ttacatata ttggctaaaa 95760

aaaattacat atataattga ttgcatagtt aaaatgagta attgatccaa aaccgatcaa 95820

aacctttgat tataaagggg aaaaaaaacc ttgttgatgt tgatagggaa acttctcagt 95880

tctcattgtc aatcagaact tgcaagttgc atttccacaa actgagttga taatttttgc 95940

tttcaacttt aatcgggcca gtttcgtgac cgtgtttgag gagcacagca acaccaaaca 96000

tacttgccct tgagggtcttt accctcaagc caatgatagt aactccaaag tcaaaaagtt 96060

cagtattttt ttttttttct attttttctg cagcttgaat gacacttga tttagacaaa 96120

ggttttttca taataattat aataataata attttcaagt aaggttgagc ataggacgag 96180

ttaattcaaa caattttatg tcattatttg acctgaaatt ttaataaaac tattttattg 96240

agtcaaaacc atccaactac ttgccttgca cccaataaaa aatgagaaaa aaattttact 96300

aagaaaaagg tgaggaaaaa gtatcagaaa tcataattaa ataacaaaaa aaagtatatt 96360

aacaaaaaaa gtataagaat tttatccttg aggtttgaat attttttcta ggtgttttca 96420

tactttcaaa tcactttgta agaactcttg catgacttca ggtaaaaaa atatagtatt 96480

aataaaaaaa agtataagaa ttttatcctt ttgaaacaa ttaacttgat ccaacttaaa 96540

ttaattgaag tataaaatac ttttaaatc gggtcatctt ttatttaaag taaaagataa 96600

cgtataagac tcaactgttt gtcaactaac caatgttttt ttcttttaca aaataaaatg 96660

ttcaagttat cgtatcatat caaatcagga aagataaaag gctcacatg gattaacata 96720

aaaagaaggg gacaagagat aagggtgggt tgtattgaaa cccttgcgca tgtgagaaat 96780  
 taccocattga ggcaaccata taatcagtaa gtgaactggt atttatggag aaaaaaagta 96840  
 gagtaattgc ttatagtgtc ttaatatctc gaaacgtgtc atgtagagta attgttataa 96900  
 attttttataa ttaaaaaaaa tgtttttatt gttataaac caaacaataa atgattccat 96960  
 taaactgtat tattattttt gtattagata aatgattatt ctccctaaaa acattaaatg 97020  
 ctcttttata atattttttt atttaaaaa aataaataat ctcaacaagt tatagttatt 97080  
 ctaattaatt taccataaaa tcaaatgaaa ttatatagat attttaatat ttatccaatt 97140  
 ataaatataa ttattttttt atagatcaca attatttttt caccctttga gtattcgata 97200  
 tacagttaca tactcgagat ttgaattaga aattactaat taagttaaac taattttgat 97260  
 gtagtgtatt tatgcaactg attgtacaaa tataattggt tttaaatatt gtgcaaatgg 97320  
 gttttgggga ttccacttac ttaatatgat ttagaacat tattattatt atttctgtaa 97380  
 aaaaaagaa atattattat tattacaatt tcttcactct tcttatatac ttttttaaaa 97440  
 acgaaaaaaa aaaatatga gtttcgttta tgattaagaa aaatgtttca taattaaaaa 97500  
 taaaaacta gtgtctgcat catcacatta ctaaactagg ttgcctgtag tcactactga 97560  
 taaaatcaaa ggaagattca ctgacgcaag aggaagtcac catagaagca gctataaaaa 97620  
 cacattttag ttccagcat gcaaaacttt tgattacgag agacatatct ccatttataa 97680  
 gtgtcttgaa caaattatga aagtagttag caacttagca tcctctctct ctttttttct 97740  
 ttcttaatat gcaatgttga tgtcccttac ttctcttttt cttgtttttt aaaatacaca 97800  
 acttgggggt gctactgttt gggaagctag agggagtgtt gcaggaaaa atattttaat 97860  
 ttttcagaaa cttggaaaac ttaaatagta ttgatattga gtgtcttttt tgttttcatt 97920  
 ttcaagaaa acataaaaa acatgttttg tttaaaattt agaaaatat ttcaaaaaca 97980  
 tgaccaattg aagacaataa attagtgtct tcagtttttc ctgaagggtt gaaaatatga 98040  
 tggatatctg gaatatataa aattaataaa taaacattca aatcaaatgt tatcttaatt 98100  
 ttctgttttg tatgtatttt ttaaaaaata cattcttgag aaatgatatt tctaaaaaaa 98160  
 aattttaata attttccata aaaacttttt cgtgaaaaga ggaggttaatt ttactttttt 98220  
 tatttttaatt ttactacag taaatatatc atattttttt attaaattat ataatatgat 98280  
 aaataattaa aataatcgat aaaaatatc ataattattt aatttttaaa aaaaattatt 98340  
 aaaaatttg aatggtcaga aaaaaaata tttattaaca tgaatatga gcatgaagat 98400  
 tcattatttc caaaattaa aaaaaaaaa aaactcctac aataccaate gtcggttaaa 98460

ttcttccata gcttgagag ccccaatagc cattcttagt gacaaccac ttggattgt 98520  
 catttggat aaaaccagca ataggaaaag ccaattgggt gtgtttgctt gacgtatcac 98580  
 ttgccattat taattattca aaccaacaca aaatagggaa tatgaaaaa aaaatcactc 98640  
 ttgtgtcaac cgaccattcc aaatgcaccc accaccatca ttgtagcttt atacacatca 98700  
 ttgacaccga gacctaatag cctgtaatta tttaaactta tccaatcttt ttttaaaact 98760  
 atttcttaaa tgtcttcttt cctcaagcac ataatccaat ccatatccaa ttttatctat 98820  
 tgcttatcgt ttggccttt tttcttttac aattacaaca atttcgcctt ttttcttatt 98880  
 agccttacgc agttggtgac gccattgtac cttgcaatcg cagaggtga aggtgtccca 98940  
 atgtccctca caagtacaat ggtgcgccct tatgcttcatt tegtgtctgt gcctttttcc 99000  
 gatacacgta tggttccttc ctagcaata ttttatggat taaataaagg ataattgtct 99060  
 tgtaactaag tcgaatctaa ggtttaacaa atttaattca ataacttta attgagataa 99120  
 tcataatttg tcttttcagt tatatttttg ggattcttat gaatatgac tctcttttaa 99180  
 ttttataatg attgtcttta ttttatactc tcaattaatc cttaatatga gtgttttata 99240  
 atgacttttg aacgatttta tagtttcgga aaatattata tacaataata atttatcatt 99300  
 attaataaagg aattgattaa aattgggtga cagagattaa atcatgggag aaaaatctta 99360  
 aaagagtga caaaaaata atgcacaaca tagtcctca gtcttagtgc acaatgagat 99420  
 ttgaagtcat gtaacgatca acaataaaca agaataatat aatcaatcat gtataataat 99480  
 aataataata atagaaatca agaacaagta tgtccaaaaa atttagggag ggtaatcgaa 99540  
 gattcctgac tcaatgtagg ccatattgtc taaagagtgc ttccaaattt ccattgaaat 99600  
 taaagctcct ttttaggcac ctagacctcc taaagggtag ttaggtggtc tcggtatttc 99660  
 actccacctt gttaaatttt ctttgtggga tctcaacgta cacattcaaa gtgattgaca 99720  
 agttttttct tccagaattt tagtaattgt caaaacaata ttgttggtct gataggtcct 99780  
 caaataaatg gaccagagaa atttttattt gaatcataat aggtaaattt aaaaatattt 99840  
 attaatattt ttatgtatct tttaaaaaatt attaaaaaag tttaatctat ccataattta 99900  
 tttcataaaa aatataatcc attcattagc attttttacg gatatatatt cagcatatcc 99960  
 ataattttat tttaaaaaat taaatattat ttgtataaat tttttgttta attcacttta 100020  
 aagaagccat tatttattta ttttttaaga taatttttta aaaataaata tttatttctt 100080  
 aaatttaata tatttttcatt aatatataag aagatattgt atttttttaa aaaaaggaa 100140  
 tccaaatatg aaaaaatatg ttggatagct ttttatctgt aaatagctta ttataggtgc 100200

catgagattt gaaaataaaa gagtataaat tggtattcat cttttcaatt atatgaacca 100260  
ttcaataatg aaggctacgt tgggagagtt aatcataact tctaccatgg gaatgttttt 100320  
cttgtttgca agatatcatg gttttaaaat cataacttaa agtgaagaa attaaagag 100380  
tctaagtgaag aagaaaaatg taggaagaaa tagattacaa taatatggat aagtttaatt 100440  
tggattagaa aagtaaggga gaaagagact gcataaatgt aaaagattta agtttgaga 100500  
aaaaaaatag gggaagaa accacaaaaa tgtaaaaaga tttgagttta gaaaagaaga 100560  
tattggaaaa atgtaaatgg ttggaaaata aagtgtttaa atcgcaata caagaatga 100620  
aaagacatta atgataaaaa aaataattaa agtattgaaa taacaaaaaa attaattaac 100680  
aaactacaa ttaattttta tcaataaaaa tattaaaaa aaaatgttaa aattgaaaaa 100740  
aattaattat tatgtatga caatataaaa taattttctca ctgtcataca atcataattc 100800  
atgatattga tgactttaag ataattttgt taaaattaat aaacttatga taaataataa 100860  
tctataattg gataataata taaaattatt ttaaatagtt attgcataat ttttttttac 100920  
aattaatatg gttgaaaaca ttgtaacaac ctgtcaatat gagaacatta gaaacttagg 100980  
atatgtaaga tatttacaat tcacaatcct tattacatga tgcaacttag tattttttaga 101040  
cagctatttc taactataaa taacactata gacttttttc aataactttc aaactcggtc 101100  
caatcagtaa ttcgatttaa gtagtgggtt actggtccaa tgaactagta attgacttaa 101160  
tataatcaaa aaaattttaa attatatatg tattttatat atataactat ataactaaca 101220  
ttatttgatt aagaaaaaat cattttact tcaaaaataa aattgaagta ttaaagttga 101280  
taacaagtcc acaataata atttaataac acaattacac aagtttcaga tttttttaga 101340  
gcatttttga ttttttatac acgaaacagg ttttaaaaat aaaatcggtc cgtcgggttt 101400  
accccaaatc ggtctggtct gaccggatca tccgagacca gatgtatgc cggtccaata 101460  
actaaatcaa tccaatccaa atcgagacaa actggatttt tgaatatgg tttgaagtgt 101520  
atttaaaatt ttatatact ttttatatct atcatgttat agatataaaa tattgataag 101580  
ttttattaat aattaattat tataaaaaaa tgatgggtcta tcaatcacia tttttttttt 101640  
cattacaga agttaaaact gagattttaa ttaagaaatc taagttaaatt ttactcaat 101700  
ctaatacat gttagtaaaa ttttatacag agtctaataa ttatagataa attagaattt 101760  
taaaatatac acaattaatt gccataaaa tatctatttt tcttctcaag aaaattacga 101820  
gcatgtttgg atatttgttc aaaataggat acacaaattt caagataaat ttaatgattc 101880  
agaaacttta ttttgcaaaa tgaaggattt gagtactttc gtaaacctca ttttagccta 101940

aaggtctggt tgaacaataa aacaacatt tactacctta ttattgcctt tccttaagta 102000  
gcaagttatt aattccttta ttttgcaagt gaatacgttt catattgcat gttttgacct 102060  
ggactaatgg acttctccat attgaccaa ctcttaaaaa caatctcgat tgtcttcttg 102120  
cgaaattgat ttccatcgat ttctctctat tttttctttt ctaaaattata ttgacaacta 102180  
tatatatata tatatatata tatatatatt gtaacaaatt ttatgatgca gttagataac 102240  
aagaatcaaa caataaaaa atataaaaa aagcttaagt aattaatct cctaaaaatta 102300  
aatcagataa attaaaaatg tttgtatat gcttgattgt tgacttgttg gaaaaaaatt 102360  
taaaagtgtg ttggctctcc tatcttgtca ctaggaaaa ttaagcaaca aagagaacaa 102420  
gtaaaagaaa atgggcaag aaaaaagtgt gaattgttg ttggggaatt ttgggctggt 102480  
cggctagaac agatattgct agaacaaaa agagtaccaa agtagatttg gattccgaag 102540  
ttaggaaaga agggttattt taaagtgcca ggtaaatgta aactccaat aaattaaagg 102600  
aaagggaaag aaccttacta tggaaatgct tcaccgtgac attgtatgta ttgggatgaa 102660  
ccttcataaa attaatctta aaagaattga ttttataaaa ttaattttga agtgagatag 102720  
tttgtgttta aatgttttat tataaaatca agttaaaaa aaaattcagt acaaaaattt 102780  
tatatcaaaa atagaagttg ttgaaagtta ctttaacaaa aaatcaatta tacgtttcaa 102840  
acataaaatt aaacatctat gtatctaaaa tcaattttgt attcagaatc aattttccaa 102900  
agttaaatca aacatgcact atttttttta gacgttatta tataaatcta tataaccttc 102960  
tgggagtagt ttaaacaaat ctaatctgga ccaacaagg actcaaatta tcttccgggt 103020  
gtgaagcagt acttcatact atcaacaaa tgatagacat cgcatttcct ttacttttat 103080  
tattagtaag aatttctatt ggcacttttt cctttttcag ctctttttta gataaagttt 103140  
gtaaaagatt gttcaaccaa aaatcataat gaaaacttct gtgggtgtgg gtctacaatc 103200  
ctgcaaattc acggaagcaa caacattttt atcactttca actataatat ttgcatatat 103260  
aagggcattt tcgcattacc caaatttctt tctttttttg ttgttgtttt ggttttgggt 103320  
tacgtttggt taatacaagt atagtataca ccgatggat ctgcacaaa aaaataggaa 103380  
gtcatatccc ataagaagg aactgaaaa tgtgtgatga tcgtcaaat cgecccaaca 103440  
tcatctgaaa ggccaatat tgacacaaaa taggtctgtg gtttgggtg attgctaatt 103500  
aaaatatcaa accagaagac aaaagggaca aagaatgaca ttagataaac aacttagcga 103560  
gtcagttttg aacccttttg tctcaatggc gctgcatalc gtagcattaa ttaattgtga 103620  
ttttaatata tctacacgc aacatcgatc atatagattc taactggtt gggtatcttt 103680

acatgattttt ttgtcgaaat aagattatgc aaacaaatat agatggagaa gtactatata 103740  
ttgtataaga aaatatttat gagggaagtg taaggatggt tccatgaaaa aaagaagatg 103800  
tgggcctgac acaactcctt catttccttat atgattatgt catttaaaaa gaataataaa 103860  
tatcatatta cattttattaa caaatatttt aacacatat aaaagtggaa ggattgggtg 103920  
agaggggtttt tcttatttat aataagtatc atttttatgc taaaatgttc tttttttttt 103980  
ttttgcgtgt gcttttagtc agctctctct ctctctcaaa ttccataaaa ttgtggaatt 104040  
agtcaatttc gttttaatta ttatattgcc cttttcaatt tttataaat tcattacatt 104100  
gcatgtaatt tatgtaaata tttaaatata ataataact taattgtata aaaagtaatt 104160  
gtaaatttgt atacttcatat gcaaagaatg tattaataa ttaatgcgaa ttaattagat 104220  
attagaatga taactggatt tgagcgtaat ggatacgcaa tataattact tatacgtaag 104280  
ctttttatgt aatgttatag gtaacggaat gtagtatagc actcatccag catctatatt 104340  
agccctcata cacatataat aaatataatt atagaattaa ctatttatta tatttataat 104400  
tagtttaaaa atatttttaga tattttttaa taaaataaaa cattttactt ttaatttttt 104460  
tcagttcata aaaataattt aaatgtaaaa aacattttta ttttaataat ttataaaaa 104520  
aagacacatc gaaaagaaaa tatattacat gaataaaaa ctaataatat gtttggtac 104580  
atacccatgt gagtataata taagtatgga ttgttttgg aaggcgtaat gctaagttta 104640  
acattgtttt acgcatagat atgaagttag ataataattt ttcttgattg cggatattcg 104700  
aatggatata ttacacaaat ggctagcctt actaatgatt tgacatatta ttattaagca 104760  
gtaccatttt ctactacatt ttaattactg aaagaatgat aaatgttgta agatatcaat 104820  
aaacattgac gtaaaaaaac aatagcattt taatattatg cttaggggtc acaccattta 104880  
actaatctcg ttgatatact gcttagtgct ttagatgctt tgatttatta ggcacaaatg 104940  
cttttggaat ttggtatatt tatcgtggag gtcaggtc aa ttgtatttaa agtgagagag 105000  
gatgcacga atcactagcc tagtagccca atagaagtaa tcggcatttc gaaggtctcc 105060  
attaatatcc caattcctta gctatcaaa catctaact aacatagaga cgtaaggatc 105120  
tgccaattct gatcataaca atcaacgcaa gtattcgatt tcaataataa taataataat 105180  
aataataata agatgggaat gcataggtag aaatgttcgt attaatattgt taactataaa 105240  
tttaaatggc atgaaaagca atgcataaaa ccagccaaat gttaggctgt gtgtataatg 105300  
taagacagca aaaaaatatt gaaatgatta ggttttgtcc aatctcgtgt gaaggtatga 105360  
cgtcccttc tactctaaac actccacacc accctcactt aatcatataa aactcataat 105420

ttcttgactt tctaataattg ttggtggtgc cgtgtagtg cccgattctt atttaagtta 105480  
tcatttgttaa gtaatttcgt ggatgggtgac accaataata acttatacta acaactaatt 105540  
aagaactata gattaccaca aataatataa taaataaaaa cgacatgaa atattaaatg 105600  
atggtagctt acattattac acataatata taatactata tgatataaat gtagccttgt 105660  
caacatgtga ggtgagtggt agtggtgtaa atcttttagta ctatttgatt cttaaaaaa 105720  
aacaaaatgt aacaatgcgc gcgttgatat attccgggtg cgttgagaca gtgaagggaa 105780  
cagtgacaga gtogtggctt acctgggtgc cacattagcc atagtggaaa catgacgatg 105840  
actgogacac ttggtagtgt aaatatattt gactctctgt agctcaacac gaccaagcat 105900  
ttaacaacaa aaaaaaagt tccactctcc aataacaaga caaaaagtc tcccttattc 105960  
taatacgttt ttattttttt atttctatgt atttgtatat acataaataa ataagttaac 106020  
cccatttaca aatgggtggc tacgtttcca ttttttatta tataataaaa attgaagtta 106080  
gaggcagctt ttgaaagcgt cgtcttctga accaacacaa tgatggaatg gaagggtctt 106140  
gcgttcttct gcgggttgat agatagctag ccatgttggc atgaagattt gagcgggaaa 106200  
ataataagac aacatacata acgttttctc aaagtgggtt tttttttttt ttttttttat 106260  
cataaggttt ggctttttat ttttttattt ttcagttttt gtgataatt ctattaactg 106320  
tcaagggagg aaattggtag agatcaaagt ttggtgtgac ttgccaact aaatattatg 106380  
acatcaatag atacttagat agatagatgc ctttatttat ctttgttttg aactactatg 106440  
attactggta ctttctctcg gcagttgtta cttaaaactt ttgttgtctt tggaaaaagg 106500  
gaacaaaaaa aggtaaaagg ggggtctatt taggatttg ttttgttct ctatgtttgc 106560  
gaaaaatgat gaggtgggtt tgttgcccta ttaatgggaa aatcgtaata tagattctgt 106620  
tgcataggcg ttgtctgtgg cgggtgtcat acaaacgtgg tcgggggtgga ggaagaaact 106680  
aaaacatggt ttaattaaat gaattattog tttattttgt gcttctcttt ggcttgatg 106740  
gtctgggttt tttcttgaga tttaggttct gtttgtttat cgtggattgt tagatattaa 106800  
ctaaggaatt tcaactctog tctctctcat agtacttctt ctgggtccaa gaagggtgga 106860  
acagtagcaa ttacttgatc tactctgtta ctttttctg tctttgtctt ttgctctgc 106920  
cacccttcta ttcttagttt ctctctctta tttttccct tagttgtttt tgttattaat 106980  
ttttgaagt aatgtgttg gcttgcaate ttcaatgcat tttaagattg attgagcttt 107040  
aattcatctt tttttgggtt ttgcacagcg gcagtttgcc tgcactgac agttcttaga 107100  
gcatgaattc agcatcccc cagtttgttt cctcaagaag gatgagtggt tacgacoga 107160



tccaccagat tagtatgtgg ggagaaggct tcaaaagtaa tggaaattta agtcagcca 107220

tgccctctaatt tgatgagaca gacatgaagt ttgatagtcg ggtgcaagtc tatgaaatcc 107280

gtgctttgtt tcattggcaa ataaaaatgc ctattctgcc tactaaccaa attcgtttgt 107340

gctgtttaat acccagtcag aggatgcttc tcattggtata ctggggagag ctaataagta 107400

tgaccaagaa gctaacaacac ctactgataa ggtaagagca tgttatgcaa actcgtttcc 107460

tcttgctcta tctcttcccc aaacaaaga aaagaagaa ggcttatttc ttggtagtgt 107520

tttgtaattt acaacaacat atatttcttt accactgaga gcggtatgaa attttgtctc 107580

atccccacat gggttctaat actttctgtc agtgggtttt ccttgacttc cattagaaaa 107640

aaaaaactat atatacttg gaactcttga atgcttctgc caactgaatt gtcagctcaa 107700

cttacctgta aatgtagata tagataattt tgagcattac aatttcttgt ttctttcatt 107760

ggggcatttg tctggcatca aaatgttttt ttttccatta aatatgttgt cattttgaag 107820

cttcaagtct tattctgtag ggtgttgttg tataatgtaa cgaagtaac aatccattta 107880

gaattatgga cagagttatt atttggaatg ttttggcatg gaaaaacaa gccttttcc 107940

ttttctttcc cctccatctt gagctttatg tatacaatta gggtgacaat gacaagtatg 108000

ttagtcctgc cattttggac tttaggaat ctgtatttaa ttggatgta tggcttggat 108060

tcatactctc ttttatgcct atatatggg gaaacaatta gggatctaatt atggtcacaa 108120

atgcagatac aaagacgtct tgcacaaaaa cgagaggctg ctgtaagag tcggttgagg 108180

aaaaagggtt gaatgaatat acgttgatac attgtcagac atgcaattg tgtttagata 108240

tactaatgtt catgtggttg ctgatattgg ctacagggc tatgtgcagc agttagaatc 108300

aagtctgttg aagttagatc agttggagca agagcttgaa cgtgcaagac agcagggtatg 108360

gagaagatga gttaatatc tagcagtttc tcttgggat cacattatgt tgatcagcat 108420

cttctctctc ttccctctct cttaaatatc ttaaggcaac aaactgcag ggaatttata 108480

taggcggttg gttggattct aatcatctgg gttttgctgg atcagtaaat tcagggttgt 108540

ttgtatttat gctactgctg ctatatattt gcattgcaact tggagttttt gtgtcttgtt 108600

attccatgtc ttgtaactt gatgttttgg ttttttaaaa ttttctctg ttggtgaaac 108660

agtttttcca aatgtgctta aatctctac tattggacat ctgaagacta aagtcttctt 108720

tgtgtctgcc atctcaaaat attcagaaac attttctcta gtctgtgcc aaccttttg 108780

tttcaaatca taatcagttt tgtaaatgaa tccattattt ttcttccaaa tatccagaaa 108840

gcggcagaat gtttaagacc ttgatatacg ccagttggag aattatttat atgtatcttt 108900

catgatacc gctgccagc tagattatca ggtctcatc attaaaattg ctattccttt 108960  
tctttgttgg atgttgtaca aaaccataat taaactgttg gtaagttgag aaaatattgt 109020  
gtaccataaa catatcttga tagtgtctta tattagcccc cagaacaac ttttgggaat 109080  
cttattggca acacaagtat tgaattgttt agaatttttc atgggtgccat gtaagcttaa 109140  
gttagtttgt tttgattgat gtaatatgtt tcttggtttt ccatgagtat tttttcttct 109200  
tgtgttatac tctattttcc aatttgttct tgcattataa taataccata ttgacaggaa 109260  
ttacaacctt tgagatggag tatggacatt ggggtgaatg gcaaaataga caaatcacag 109320  
aactgagaaa tgctttaaat gctcatattg gtgacgtaga gctcaggatt ttagtagatg 109380  
gtatgatgag ccactatgcc gaaatgttcc gcatgaaatc tgctgtgcga aaagcagatg 109440  
tcttctatgt tatgtccggc atgtggaaaa caacagctga aagatttttc ctatggattg 109500  
gaggctttcg cccctctgag cttctaaagg tgatgcctca tatttaataa acgaaagtag 109560  
ttaaaaaagt tattcataga aatgtagaat gtgtagatct aaattcagtc attgtctttg 109620  
aggactttct cattattatt ggattttaat caataacttc ttctactgtc ttcttgagag 109680  
agagagatgg gggaaatttg aaactgcatt catttattac agttttaata cattatgcaa 109740  
gatttactca aatgcttttt gtacttaaaa tggctatcac tcagtgtcag tgatagtccc 109800  
agagagctgg gtcacgcaag aggtgatatt ttagatttag tcttgacttc ttgcttttgg 109860  
aaagaattgc atcctcacac ctgaatttca aattttaaaa aagggattta ttgagatact 109920  
agattaattt gagattattt ctattgccat attcaactat ggcagtctgt acgcttghaa 109980  
tgtagtatgg tgcttgctga caactttatt ttaagtgcag gttcttggcc ccttgattga 110040  
accctgagc gagcaacaac gtttggatat ctataacctt ggacaatcat gtcaacaagc 110100  
agaagatgcc ctttcacaag gtatggacaa actccggcaa acacttgctg atagtgtagc 110160  
agctggcgag ttcattggaag gaacatata tccacagatg acttctgcaa tggacaagtt 110220  
gaaagctctg gtgagctttg tgaaccaggc agatatcctg ttttcttgat agacattagt 110280  
tctcaattta gtaaaaaatt ccttttcaca acttgccata tcaatcaatg gattcctcct 110340  
gccagttgca atttttatcc agatgttctg tgaaaaaaa tcatgtatgt gttctggtaa 110400  
atatacttcc taatccacat gtgcttattt gtgtgtctgt ctctctgtct ctactcacta 110460  
tagattcaaa atagaaaagc aacagtttag tgtgctttaa atatcatgat tctatgtcat 110520  
atgaaaatgg aacagatgtt tggttatttt aaatgtatag ttttttgggt taagggatatt 110580  
gttgaagtta tttgaagcta ttgatcttgt aagtcaacaa atttttcata tctaagtag 110640

ttcttataca tcattctata aattgttttg tagtgcaatt gattttaaat ttactttttg 110700  
gatgtgtaaa acccataaatt tgatttgaag tagatgaaa atatttcata aaataaatc 110760  
cattctctaaa ttcttagcac agatactgaa actgaaatat ttttcattct aaagaatcct 110820  
cagtagaagt ggaatacaac aatttaatca atcaattcaa agttttaaaa cataattagt 110880  
ggtttatgat ctattttaag catcctctta cgatttgggg gcaagtcatt gcgtaggca 110940  
tttaatttca tctattatat tagctgtgac ttgtggttca cttcaaaacc ttaccatt 111000  
catcttggtc actgatagtg cttgctagcc agatttttat ttagcgttac caacttcca 111060  
tttagcattt aggtagattc acttaggat cctatatcgt gtaaatgcta aggtcataaa 111120  
gtacagtgat ggttaattga tgatacatta acagaatcat ttttgaagat cattaactgc 111180  
tatatacact acaaaattac aacttcaaaa agttatagct tggctaata ctattagaga 111240  
tttaaccatg ctctgtgtga aacaggctga tcatcttcgg caggaaactt tgagcaaat 111300  
gtctcgaata cttacaatcc gtcaggcagc tcgatgcttg cttgcttttg gggaatattt 111360  
ccaacgcctg agggcttga gctcactgtg gtctaataga cctcgtgaac gtcttagtc 111420  
tcacaacaag gttacaagag cgaaactca taacatgaac gtgatgtgat gatgaatatg 111480  
ggtcagtttc ttggccaag agcatcaaaa ccatgtatca tgtgaggtcc aggttgctca 111540  
catttttctc gcagttttgc gttatggaga agttgtaaat atgactgcta gtcatttta 111600  
tgcatacag ttgtaaata aattaatttt agacgcattg tttatatcgg attgttgcaa 111660  
aaatgaaata tttagaatct gccatgagtc tgttggaag ttggttggtt agaattcgta 111720  
taatgttatg ttgtgtgttt ttgaaatga atcgactcta gactctctta gtggacataa 111780  
tttatttttt gtgattgcta agggtagctt gttaaaaatc tgtaagtgtt tattagtgtg 111840  
tttagaaagt gttaactcta ataagctcta taaatagagc attcgtagtc ttattgtatc 111900  
ttttattcaa tcaactgcaa catgatttta tatgatactc ttcttctcta ctctgatctt 111960  
tatggtgtat ggccaataa tgttggtgtt acttgtctct ttctcgaag ttctgatagt 112020  
gatagtgtct aacaatattt atcataaatt tcatttttct tggatggata cttgagattc 112080  
tgtcttagat gctgagttag aaaatattaa tgtgctgaag gttttacgtt ttactttgct 112140  
ttcttatata tatacttata ttggtaaacc aaacatacac accaatccac ttattcacat 112200  
tcaatttata ttgtcaattg ttacaggat ttctactatg caatgggatg tgcagctatc 112260  
aggatttgtt tccttatagc aaattgggtt gatgatccac ttgtagctag caggatatat 112320  
aatcatcttg attccagaca gaattctaca acttagatat agcacagtcc aaagcatcta 112380

caaaaaatct gcacatcagg aagccagtca gtgaacttca acaagaatc tgaactccac 112440  
 aacttggttg aagagagcca ggacagaatt atttatatac ctgcatcatc ttgtgtgaaa 112500  
 cacattggtg gcttgtactc aaaaacattt ccatgcagcc ctctttccc aactagaata 112560  
 ccgagctcta aaatttatag gagtaaacag ttaccatgtt ttgtagaaat ttgggaaaag 112620  
 aaagcttcat atgcaatttt ttttattcaa ccaaattttc tacogctgaa tctttcatgc 112680  
 aaagctgttg ttccagcttt tgcagggtgc ttctattttc tatcggtgac aaattctatc 112740  
 cccaccatta agcctcttcc tcttacatct ccaatgactg gccaaatata taaaataat 112800  
 caacaaattg tggcaaatgt atgctctcta taatgaagca tcatatagat acctaaagaa 112860  
 aggcaaatgt atcgtaaaat aaaaatatat caaaaatga aaacatgaat atttctggga 112920  
 tataagattc aagataacca tactgtcatg tacttgcatc atagatctta aacgtgaat 112980  
 cgagtggaaa ccaacatcag cacaatgaca ttgacgctc tccttgtaa gaactctgag 113040  
 cacagcaagt cccccagca aacatacagg gttctctcca aaagtattaa attgaagctt 113100  
 ctgggcatc acgcttgcta tttctggagt tgtaactaca gctgctaagt gcaaacatt 113160  
 gccaatacc tgatgcatgt taacatgagg tacaatcat acaatggaat tctagaaaag 113220  
 gtacatgaga ggcaactcaa caaaattggc cctatctgaa ggtaatgcta gaacaaaaa 113280  
 catagtatag tgaaggaag caaacaaaa aaaggaagct aaaccgata aaaggataaa 113340  
 cattgaagtt atgcttaaga gttatacagt gagagtgc aaagtgcga gaattcattt 113400  
 aatattcaac tatataatgt ttaagattgc ttacttaaa cctatatatta taattgtatc 113460  
 atcaaattag ttcactaat gggaacatgc tctcatgta catgcttct atatgcataa 113520  
 ataataatga acctcacaac agactgatgc tgcaactcaa gtgcattatt aagggtgcat 113580  
 tgtatcttgc aactagttaa aacttagaaa gaagagatga gggctagac ctttgccatg 113640  
 gtaactatat caggaatgac accctgtgtc tcaaatcccc aaaaatggct tcctgtgggc 113700  
 gcaaacccac attgcaactc atcagctatg cagactctc cagctttgtg tataatgtca 113760  
 taggctagtt taaggatccc gggtgccagt tcaactgtc ctctgtccc tagaagaagc 113820  
 catgacatga gaaaaaagta caaaggaaaa ccacactatg gtgtgctat aaatattaaa 113880  
 taccactctg aagtttagaa attaaacatg aaggatggat tatgcatgcc tgaatagttt 113940  
 cagccataaa tccagcaact ctctctgaag ggccatagtc aatatgatc tgcaattctc 114000  
 tggcgctac attggcatct gagccaaaa ctccgcggtg tggatcagga ttcatgacat 114060  
 gatgaacttc tccttaag agttaaatgt tgtaaaaaa aaatacaatt agcacttccc 114120

tatctttagt aaaaagttaa ttcacgttc acaatgatat cgttacaact agagtgcac 114180

ggattatgtt acaatgatat agtacatgtt tgtgtatgtt ttggatata tatatgcacg 114240

tgtttgtgtg tgtggacaaa acctctggta ttgggtatct ctacgtgttc atagccgtgc 114300

gtccaattgt actagaactc tctccatgat atgcattcct taaggaaatc ataccatata 114360

taccagtgtg taaacgggcc atcagcatcg ccaagtcatt tgccctgtga ccagaattta 114420

aaaaataaac aacctgtaat ataaattgct tagttgaaca ccatgttcca tattgaatga 114480

atgaaatcta gcacactata aaactgaaca ttgcttttca gcaatctcaa atttggcatg 114540

tcacagtttt acacactgtc atataatcac aaattgatta ttggaattat tctaaaaata 114600

ttatttttaa agtcaacaaa cttaccatat atgatgaatt atgattgaat gtctaagtaa 114660

aattttttag tgttagtgtt taacctatct ctcttaccat atttatcatt ctctatagca 114720

ttaaattggc agaggggtta ttaaaatttg taaatattta tgtaaagaga aatagatata 114780

aaatattaca aaattttcat gggactagat catgtgaaaa attaatgtag gaggatcatt 114840

cccttgtttg gatagtctaa ctgaaacaca aattatgatg ataaaattga ccttaagggtt 114900

tccaggcatt ttggatgcca aggcctcgcc aaaatcagct attgcattgt gtatataat 114960

ggttgttgta tgcgtcaaga gtttactttg ctccattata gcattcaaaa ctgcaggatg 115020

gcaatgtcca caagaaatag taactattcc agcaaaagca tcaagataac gcctgccatt 115080

ctcatcaaat agatgttgca tcatgccctc tacaatgtta agctatata aaaaaaaaaa 115140

ttgttaaaaa ttacaaaaat atgtggcact cttttcatat ttaataaatt ttatttataa 115200

ttttgtaatt ctcaataaat tttaattggg tgtcagacaa aatatattaa agaagatgtg 115260

agctataact cctccttact tatcggtatc gatgatggga ttaacatag acagtctctg 115320

tggagaccca ttttcaaata acgtaactat ttttacttca tttctttgac tcaattttca 115380

gagaaaaata aactaagaaa acatagagaa aaggtttagc acttatcata atatagtaac 115440

tgataaaaga ttgattaatt tgggagactc acaagttttt gatagaagtg gaacaaggaa 115500

ggaccaagga actttttccg ttgggcaaaa acttcatcag ccaagggtcc ttgttatggt 115560

cgtggcttgt aatcaaatgg tggaagctga agcctcogag caatatcatg atcacttgtt 115620

gcagttagg tgctcaaaca agaagaagaa ccaaaaagcc tcttacaaac accaatgctt 115680

ctctcacca tcgccatgtc taaaccacct ttactaaac actacctata ctttatctc 115740

aaactcgatc atccattttg taatgctttg catgccatgt ctacagaata attgacatgg 115800

aatcttgtaa atttggatta totaacattc aaaatatatt atgcggtaat tcattttatt 115860

caagtaatgt aataaatgat aatataaagt ttaatttttc aagtcaccta ttagatataat 115920  
attgctttcc atgattgagc tagagcaaaa aagaaaaag tagactgaat aatgtggttt 115980  
tattgccttt tattcccac ttgtttgctt ggtgtaatta caccgtccat catcgcat 116040  
tataatgaaa agaacttaaa tacgatagta tatgatcata tatatcatat aatgagtata 116100  
agtgcattg aaattgatcg cccaactaca accatcgctc tgatcatggc agacatataa 116160  
atgtacatcg attgagtcac caagagataa atcctctttt attttgattg agtttatagg 116220  
aattaatcat ttattacttc tatttgtgta ataccactc cagtataaag taattcagac 116280  
cttcattcatt tttcatgtaa ttgtattggt gtaggaagct ttcaaggga gaagagagag 116340  
caaagtgatg gaattaacta ctggaatagg aaagtgtgta catcaagcaa gctatgcacg 116400  
tacagcaaaa ctacatgaaa acaagagggg tagttgtaga gacttgagat aaccagaaga 116460  
aaatttgtga gatttttcat tccattattt taatttcaat acaattagca tttagcagta 116520  
cgagaatcat gtaaggaaa acttgaatag acgagacaaa gcagggtagt agttgtactt 116580  
gtagatactt gagataacca gaagccaaat tgtgtagatt ttttatcccc attaataaaa 116640  
tacagcaagc agtataagaa tcatgcaaag gagaaaactt gcatagatga gatgaagata 116700  
tcacaacatc aatgaattat ctcaaaagcc tgcaggaaaa ccaccctttc ttgggtcact 116760  
cactgccaca agctttccat ttccaatccc tttattttgt ctggacggca cagaattgtc 116820  
tagaacaata aattggcaaa tagtcccacc ggcaaggccc tttaggacat gacccttact 116880  
tttaagggtc tcctgatat cagcaggaag ttcaaagtgt tcacctaaag tgggtccaatt 116940  
ctcataatta accacgttag gtagtagctg cataaaacaa catggcgagt ttctaatta 117000  
attagtgatg gacaagtga gataagtatg cagtcacaat caaaaagag aaaaatatca 117060  
gaataacagt agcaacataa agacaaaatt cttattactt aaaatgaacc aggagaaaca 117120  
gtttctgggg acaaaaatat ggattcaagt tagaagtttt aaattctgat aagttataac 117180  
atttttctct ggttttaaat ctgtcattca ttatatatct ataactgtct ctgtgttttg 117240  
aagtatttat tacagtttat cttgtgacat aagtatttgt gcaagtattt gggaaaaact 117300  
atcaaaaatg cgtatgattt tgtctataag ttgttttag cttaattcaa tttgttctct 117360  
agtataattt atgaaaacaa tttaacaact gtgaaaacag tttaatacata tttcctcttg 117420  
aattatagaa acaacttata cataatcact taaatggcaa acacttgttc aataagcaca 117480  
taattaagtt gtttattagt ttgtgcatac cagagcttac caatcagagg cagaaaaaac 117540  
aagcaacccc ccaaaaaatt taaatacaat acctgatgat agacccttgg agctgttaca 117600

gaagaaaaag gatccattcc ttttaacgaag tgattcaaaa gaactttctgc agttccacca 117660  
atgatgaaag cacctccact tgcacctact acagctttta gcttcccatc ctgtataggc 117720  
atggtaggaa atcattgcaa gtaaagaata tgagtacctt gaattaataa tagtgaaagg 117780  
atttcataac cattgcaatc aagatttggt cttttgtatg cttgacaaga gtgattgtgt 117840  
catttaataa aggttaaatt actcagttgg tccctagaat atctgaacat tttcaagtag 117900  
gtctctaaac tatttgtttt ttaattaagt ccttaacta aaatttggat ccttggatta 117960  
ttttatggca gcttgaaatg gccagaaaat gataatttat gacggtttta aaaccctagt 118020  
gactcactta aaaaatcaa tgatcaagaa ccaattaat aaattttagt ttaaagacct 118080  
acttgaattt tctcaaatag ttcagggacc aattgactaa ttccacctt tcttaataaa 118140  
ttaaaattct tatatctcta tctcttgatg aagctcaact tgagcagtag tgggtgacat 118200  
ggatgataat gtggaagagg atagccacat ttaagaaagt caaatcttga taaaggattg 118260  
attaccttaa gggcaatagt gggtgacatg gatgatagtg gccgctttcc tggcatgatg 118320  
aaattggcag gagctggttg tggaacatcc ttggtaacat ttctaggcat ggaaaaatca 118380  
tccatttcat tgttctagtc tatctctgta cttggtgaaa ggatccttga accaaaaatat 118440  
gcattcacag tgctagtcat ggaaatggca tttctctcag gatctattat acttaaatga 118500  
cttgtaccat gatcatggat ctgattccac ctgaaaacac aattcagcac aactgagaat 118560  
tctgtattgc cacaaactta attttagttt agaacttatt tttacttaga tctttatata 118620  
tatatatata tacaacttag cgcccataca tacaatcatt ttaaacttcc ttcttatcat 118680  
ctgttgcatt gttaacacag agtaacacag aatgtaatgt gaggtgcta ctatacgtg 118740  
gttttgaag attgaattca ttacttcact tgacgggttaa atgacatttg taggtctaac 118800  
tgaataagcc tagacaatag atagttatc ttgtgtaaac agtacatago tgcacacct 118860  
ctcagtttct gttagttag ccaaatgcac gtatatagga caaatatata tatatatata 118920  
ttagctaaat tgaatcagg gctacattta aggacaaata tatatatata tatatatata 118980  
tatatatata tatatatata gatagataga tagatagata gatagataga tagatagata 119040  
gatagacatg tacagcataa tcaagttgac caaatgccta gaacacctaa agtcaacctt 119100  
tatatacaag ttgcagacac ttctcctaatc aagatatatc tttctatgtt tcaagttgac 119160  
tatatcatga aaccaattgg tcaagaaatt gatattcttc aactctagct tatggagatg 119220  
cgcaagttat gatgcattaa ttgtcaatat tttctgcatt ttattcttgt cattcacacc 119280  
aaaagctgca tattgcaagt gtgtgaggta gaagacaac atggaatcta aagaacctca 119340

ttaaacttgca caggcatgaa gctacgtgaa aattagtgt aactcaaga acttaggagc 119400  
ttacctgctg ccgtaatgac tggggccaaa agttttatta tcatttatgt catttttcaa 119460  
cacgttagca aaacctatgag aaagcatac agaaagaacc ccagttatat ttacaaaatc 119520  
agggtcacca agattcattc tcacggcaaa tacatgcttc aaagcttcaa ttctctgatg 119580  
gaaaccaagg gcaccagaaa ggccagaagg gagtttatat tgatctagaa tgttcagcag 119640  
ctgtaacaaa agacagcatt ggctaaatcc ctttaattgc taattataag atagatataa 119700  
agaagaaaaa aaattattaa caagaacaca aaggtagtga atcagataga cttagtagca 119760  
tcattgggatg cccacctgaa gggggaggca tgccaaggag ctttagacca agaacatcat 119820  
tagatatggg ttctctctgt ttaacagtg aactttttag gtctttcatg ctcaatatcc 119880  
ctccagcatt ctggacatct ttaactaaat taagaccaat caatccttca taaaaggctt 119940  
ttggaccaga ctctgatatt gttctgagag tcttggttag ttctttattg tagcaaatgc 120000  
cacctatctt taagagtttt ccatttgggg caaatatgct acgaagacc ttgtcttcca 120060  
atatacttga ctctgtttct tccatctgct tgtggaggta tggatgatac ttaaaccttc 120120  
tacgtgctag aatctcagct ggcttttcaa gtcttttcca tgggaagctc ccattgttgt 120180  
tccacgcttc atgaaggcca gcaagtctc caggaactgc tacagatagg ccaccttttg 120240  
ccttcaaagt agtatttcca gcatacatat cctgaaaaaa atggataaat catgtttaat 120300  
acaccttacc aattgtttat aattacttta aaaaaattaa gtactacata ctttttatgg 120360  
tttaaaattt agttttgtta acatgggtgc tgattctaaa catccaaaag taacaggatt 120420  
ttattaagaa agaaaacatt tatcacactt atgtttctaa tgcaatacaa actgtctttc 120480  
ctaataaaat tagtctcaga aactattcct cgagataaaa agattgattt aaggggttca 120540  
catcttccaa tatgtgttct agcatacaca aagatcgtag gatcaaacct ctaaccacat 120600  
aatcaagggt tatgtttatt taaccaacct gttgggttaa cattttcttt ttaatacaac 120660  
aagtacttat ttacttctgc attaatcttc tgttttattt tatgttttta cttttaattt 120720  
ataaaactat aacattattt taaatccatt ttctgttttc aaatgtttgt acagaaaaaa 120780  
gaaaaaaagt tgttttcatg gttcccttcc ttaaaaatct ctgaaacaag aaccaattt 120840  
aaaacaatgt atcagttttt taaactaaaa gtgaagcag aaaatgaaaa cacagaaata 120900  
aacatgcctt tgtattttta tcatatactt ttatttttat ttctttaatc attgagtgtc 120960  
taaaagatac attagtcttt gtgttgaaa tatcaatgca aggatatgtg atgatgaag 121020  
aaaatgaaat aaacaaggag caaaattttc agttattttt ggttttctct ctttttact 121080



tggagtctcga tcatacaaca aatttatcaa ttgaactac agtaagcaga ggaataatcc 121140  
 ataattgaaa aggaattaat gattcagatg gttaaagagt taaaagggtta ccttggaggc 121200  
 aagagcagga gcagtttctc tcatatcaaa agcctttgcc accccattat tcaacctgag 121260  
 aagcaaaaaa gctccaccac caaggccact tgaagcagga ctcaaaccc ccaacaaaag 121320  
 ggaagaagcc actgctgcac ccactgcatg cctcctctct ctaagaacat ccttcccaat 121380  
 ccttgagcat ctacggctcat cagtggcaac agctccattg tgggcataga tcacctcacc 121440  
 tttgtgtttt ggggtgtccat ttgtttcctc aaggcttgaa gcaaaggcca aatttgaac 121500  
 caaaagcaca gcaatgacat gccacaataa cacagatata acaacacttg acatgtatgg 121560  
 ggacagaatt tgggtgtagg aatgaacat ttgctgctga cctgttgact tttegaagccg 121620  
 agacatggat ggaacttgct tcagttcagt ttgctgtaat aacataaagc aaaagaatgg 121680  
 atggcacttg gcagtgctaa aagtggaaag gactaaggca acctcggttc ttatgttact 121740  
 tttgcttttt aagcactgat gtctataaga tcctgctatt aataggggta tttgtacctt 121800  
 aagtgaaggg gtttgcctgt gttctgatcc ggtttaaaat ttaataagg tcaattttga 121860  
 ctgcacagat atagaataga taagattagg tgaatctgat ttattgatat aatgtatat 121920  
 aatgttacat ataattaat ggtagtatta aaaagtata ctaattatgt tatgttttat 121980  
 aacatatttt aatataaata atattaacat aaatatcata atattatgtt agaaaaata 122040  
 ttacttttct tctcccaaaa ttaatgttgt attttaaaaa aattatccca aaataaatgt 122100  
 catgttaatt tattcaatgt agaactaatt attttttcta gtaataacct tattaaattt 122160  
 ctttttaatt tttcaatga atactaatat tatttccttt cattttatta attggctaatt 122220  
 tttgttaaat tactattttt ttattttttt tattttttta ttagtgtta tctaaaaaa 122280  
 cactttgaga taaaaaaaaa agtatttatt aagttaagag tatgtattat taatagaat 122340  
 tatttaagaa atatactttt taaaataaaa atgtatagga ctatgcaaaa tataatatga 122400  
 attcaagcct ggtcaaatca attgaacggg ttacaccagg aacactctca cttattaagc 122460  
 atcccttaaa aagtaagtgt ttatttaaga catgtttaag ttacttaata aattggcctc 122520  
 caacttacat ttccaaagat ctctttttgt aacttatagc ttgaacttgc attggaatgt 122580  
 tcgttaccac actttcaatg atcatgcgtt tatacgttga taatgcatg gaatgaacgc 122640  
 taacaaactt ctaacgatga aacaaacatg cacttaagca tctgttctt aagtttttta 122700  
 acccaattaa tatttaaaaa ctaaaaaaat taatgttaca taaattttta gaatgtaggg 122760  
 ttatttatga gattcgttca gaaatattgt attagaaaaa attgttttaa tcataataga 122820

taaaatttag tgatgcataa tgataccatc cacttaaaat atgcaaagta aaatgctaaa 122880  
acaaagatac tccaatcatg tgatatgatt ctttgaattt taaaaaatat tctttgtaat 122940  
tgtaagtaat ctctcacgta cctataaagg ttttagaaac tacaagttta cgattatttt 123000  
tgtctattta aaatttggat taataattaa taatcaagtt tgagtggaga aaaaaataat 123060  
atggaacgct tgaataataa gcgtcaaaag acaaatggct gaagtagcac gagggaaatt 123120  
tgaagagtta aggaccactc caccaatcat gttaatgggt gagagatgct ataggaccga 123180  
actcagcctc aaccattcaa atgatgacag ttctcaatga atttcttatt taattaacaa 123240  
ttttacatgg cacccaaggg ttgttcaaaa attaaacaaa atcatttata aatcgctaaa 123300  
ttaatctaaa aaatttgcaa ttggcttaaa ttgaaatta caaaattga aaattgcata 123360  
ttttttaccg tttgattcgc ttttcaaatt ttgttaaaaa aaatcaaac caaaccacca 123420  
gcgtattata attatattaa tttttattgt gataattaat aataagtata aaataattat 123480  
gtacttttta gttgctttat aatcaaaata tataattttac tttcaaatgt taaagggaat 123540  
ttaccaatat tatgaggtag aagacagcca gttggacata agcaataaaa agataataa 123600  
aacttaacta aaatcatatt ttatcaataa tccaattcat catatattat tcaattatat 123660  
aatatttgat tgtgaataac tcatctttca aattgcaaac caagtcatac gcggacaaat 123720  
taattgctag agcatattgag tggctgcttt aagtcaaat atggcaaaa tactggaagg 123780  
tcaacacgca actgcagtc acgtgtctgt ttaattttt tctcttctca aaaaatacat 123840  
taaacctgat ttttagaact cgtttaggtc aaatggcttt ggggtcgtca agttatgac 123900  
agattgtatg ttgttaagtc ttactcttc ttaaatagac ttttcaggc ataaacagt 123960  
tatttaaaag acaattata accctctcaa gtctcaaat ttgacaatt tttttttttt 124020  
cattttgtta caagcaact attaatcttt tgaataaaa agttgacctg ggaagtgggt 124080  
gcacacagca tctaggaaaa aagtcaaaa gcaacgctgg attagaacaa aatataatta 124140  
ttgtttattt ttctttcttt ctctcttttt cattttttta tgacatcttt ttccattggt 124200  
gtaatggcag aaatcagtg tcaagtattg atgatggctc cgatctgac aaagattcaa 124260  
agacaatcct ttgtctcag ttgtgagttg ttttataaat tggacaatca aaattgcatt 124320  
acaaaattta caaactacaa tgttctcaga tgtcagtgta aatagatgat gcacagtact 124380  
actttaatta tagtagtggt caaaaatgag aaattttaag attattcaaa aatcttgaaa 124440  
tatttttaaa tgaaattaa tttaattaa tctcttgat ttttttaata ttctgtttag 124500  
atagagagaa tttcgatgct agtattgtta tcaaaagact atttgacaa gaaatcgaca 124560

aagaaagag agaaaggaga gagatgggaa gagcatagag agaaaataga gataggacta 124620

aggagaaaaa ataagtaaga agaatttcaa attcttgctt ttgactttt tagtgtaaaa 124680

attgaaatat tttatttttaa atagttaaag catttcaaga aattaatata atttcaaaa 124740

cttttttaaa acattattca aataacttat tagttttttt tctcgagac ttgagaattt 124800

tagtgacttt ttaaatctga ttttttttta aatagcgga atcagggtta aaagagtcta 124860

cggataatta gtatgaatac aaagattggg cctgagctcg aattaagccc agattacctg 124920

gggaaaacaa cagataaatt tacatgaata attttttttt tctaatttgt ctgatttctt 124980

tctgtaaaaa tattaatgc gttttatatt ttaagacat taaacgtatt tttttttatt 125040

taacattttt ataaaaatgt agaaagtttt attaagagaa tatatcagta ctcaattcac 125100

ataatgtcat tgtccaaatt ccaaatatct ggacgcatta aagtgaagtt acctgttgat 125160

tgatcatgat ctacttatct tggatgtact ataaaagccc aatatggata cacatagaag 125220

ctagcccaaa ctgttcaatc ctatcactga ccggtttaaa acaaaatgga tgagttgacc 125280

agttttggaa aagtgggtgg ttaaaaatct caattggctt cattcaaatg tgggttgatt 125340

tgaaaaaaa caaaataatt atgaaaatgt gctaaaaaaa ccaaatttta ttataattga 125400

gtttgtacgt taactactca tattttcatt ttataatctt aaaaaaaac ttgaaacat 125460

ggtgtatgtc aatatgactt atttaacta tagttttttt ttactattt aaactatggt 125520

ttgttaaaaa taaaacatag ttaataagt tatagtaag gtactcttgt tcttcaaaaa 125580

aaaaaaaaa agaggtaag atactcttgt aactaaattt tcttattctt ccgtatcact 125640

agaccatgag cattcaatga atacataaca atgatttgtt tgttggttgc ctaaaactca 125700

gacttgtacg ataaaaggat gactacaagt caacaacaaa tatagaagaa ataaaatcaa 125760

aaaaaaattg gccaaattat gttacatatg attttacttt aatttaactc cataaaaaa 125820

gttctataca tgagaagta caattgaata aaatggatg atttttttta ttttcatttg 125880

tttcaattaa tgatttaatt gattttcagt ttatttaatt atttaatggt ttaatacatt 125940

ttttttatta cagcagaag acaaaatcta tctactctc atatgcatta cccaagcgc 126000

atcagcaagt acccgaggaa gaagaccgcg tggacatgat gacagaagga ttaatacga 126060

tttgaatata ttttttagaa acaactctat caaaattatt gacggtacct aaattgattg 126120

tatatttcaa gcttaattat gtttttaata tcaactaggg gttggttttt tagtctttca 126180

aataaaaaa tatatatatt tagtctttta aattaaaaa aaaattgttt ttagctccct 126240

tccaccatct ttttccaaa caaacataga gtagttagat ttgttctatc ttgtgtcttt 126300

tagcaaatca aagagattct aaaaagagag tatttcttta attcgagaga ttaaaaatta 126360  
 caaattgtca ttgaaagac tactaattaa aagttcaaaa aataattaag ttatatattt 126420  
 attactattg cttaactttt ttccaaaaca aaatgtctaa aatgaggcta agtacacgtt 126480  
 ggtggtgaca caggaaactt cccttcatat gccttgtcag tatattaatt tacgaataac 126540  
 ttgtgtccca tgtgattcga ttttatttaa gtgtaatttt ttttcttgag gcgaagataa 126600  
 tattataata taaaaggat ccaaaacoac aatcaaggag tgatgatatt ggaccaaaa 126660  
 tgtttgttta attaaagcat acttctgtct aacttataga gttgacaagt aaattaatac 126720  
 tatcactaaa tctccacttt actatttttt aaaaataaaa tttaaattta tgatgatatt 126780  
 aattaaatta aaaactaaat aattataatt ttgatcttct tgtaattctt aatatgtaat 126840  
 tttggtattc ttatttttaa aaattataat ttttaattcaa tattcaactt tccgtatat 126900  
 tattttttat tttaaatttt cacaatttca atctacttat ttaatttatt tgagttatca 126960  
 tattataaat aaaaaatatt cattattttt ctttatatac aaacttaca agattttatt 127020  
 taagaaattt gaatccaatt ttactctac tatcaacttg tttttgtata attatttact 127080  
 ttaagaata taaaagtaca acctatgtta tattgttagt taacaataat tatcgaggat 127140  
 cacgatgatt ttgtaagaa ttatatctat gcaatttttt ctttatatac aaagctt 127197

<210> 2  
 <211> 335913  
 <212> DNA  
 <213> Glycine max

<220>  
 <221> CDS  
 <222>  
 (45163) .. (45314), (45450) .. (45509), (46941) .. (48763), (48975) .. (49573)

<223> Seq ID: 240017\_region\_G3

<400> 2

aagcttataa aaatatttca aattttatta ttgaatcaaa tatactatct atagtggtta 60  
 acaaaaatat gaatatgcat tgatttttct agagagggtg ctgggggtcca aactcgaagc 120  
 gtatttttgt aagaggaaga aaatgtcaa atatataata ttgaaatatt ttcaactaaa 180  
 ttatogcaca tcatttttac aattatttta ttgtattaca tcattttattg tttattttatt 240  
 ttcatgtaca tatttcttaa cctttttgca tttataaatt ttttaaaaaa ttataaaagt 300  
 taatocatt gtttgaatta actcaatttt tatttatgaa atagttactt aatttttagta 360

atthtgatta	gtgattattt	gaatcatcat	tcaataaatg	aatacatatt	taagtgtaat	420
aattttcatt	agttattttt	atgtatgata	tttatggatt	ttaaatttaa	aatattattt	480
atthttaccta	ttttttttat	aagccaatat	gtttttattaa	tattttattta	ccttttaagt	540
agcacaacat	ttttataaag	caaataattta	gatgaatcaa	gaataaaaaa	taaaattaat	600
ttttatcatat	tattaatata	aaaatattac	aatatcaatt	aataaaaagt	ataataaata	660
taacccttaa	agtgattatt	acggaaacaa	caaattatca	cttttatata	tctattttat	720
ttctataaatt	atattcataa	gttactacat	attattttat	cctccttaca	actactttca	780
caaatcatcc	actaatttta	tttgaaaaaa	caaaacatgg	aaaaaagtta	gacacaaaag	840
gcttggtttcc	ctagtaattt	atataagaga	ataatagtg	aattaacctg	tgtgactgca	900
tacacctttt	acttattaaa	attgggttgg	atgcaatctc	tgtctatgcc	atttcaattt	960
cgatctaacc	cttgtgtgtg	atcaattcct	cttaccacag	ttagcataag	taacatactc	1020
acctttatcc	caagctctaa	ccatttcgac	ttgttgctaa	ctagattgag	gatagtgtac	1080
tttgtcttcg	ccaacaacat	cgtccacttg	aacaaaaaat	caaaaagacc	ctgaatatcc	1140
tgaactttta	aaagccgcac	tttttatttg	ttgtgtgtga	gatttttgatt	tgatgggaaa	1200
atatcttgaa	tctgcggcga	ggctggcgga	gctatcgagg	attgtgtcat	ctgctgcgaa	1260
gcccaacagg	ccaaagagag	cgctacccag	atcgcccaac	cgtgtggcga	ctcctcgaag	1320
tgcgaaccoca	tttgggtgta	aggtcgaacc	tgcgaaaaag	atggaaacct	tggagaaga	1380
acaacaatgt	cggacacccc	tcgccaaggt	tgtctcagat	tgtcccaagc	gctgggtcca	1440
ggacaccctc	aaggaggcta	aggctgggga	cactaccatg	caggtttttg	ttgggtcaaat	1500
gtattactct	ggatagggtg	ttgcgaggga	tccccaaaag	gtcttttctt	tcttcttatt	1560
tttatcatcc	tatgtgcaat	gtaataattg	attgtacacc	ccttgtaaag	ggatagggtca	1620
tagactcata	tattattatg	cttgattctt	ttaagttatc	cctcatttga	tgcacaaatt	1680
atgtgattht	gttccctcct	tggaaattga	actgagatag	agtgacaggga	agttgggcat	1740
tttacttgct	gttagatttt	aaatcccttt	tcttggtgta	taatttttat	cagattgttg	1800
ttttttgatt	tagaagctct	ctcatattag	cttctccaaa	agctgggttt	taagtcaatt	1860
tttagcttatc	agattgttgt	ttttcccttt	attttctctt	cctataagtg	ttcatggaga	1920
agtttatcca	aatatgcctt	taaattgctt	tgggtggaagg	ggctgttagg	gcttctaaat	1980
cttgctcaat	tgtcaatcaa	gacccctcac	tgcagttaat	ttgaattgtg	taattgggtat	2040
gctattagct	gcttctcaac	ttgaggttac	tgactatact	gtgacaaact	ttgttttatg	2100

gggcatgctt ggattagtaa agcatcaagg aaccgaaatt cagtttgtaa attgtgcggt	2160
aaacggccag gtaagctcta ccttttgttg gactagtgaa ttgtggtttt atgtgtcctg	2220
attgatctct attttgtgac attcaggtta cagagcaagt gattctgatt catgtgaact	2280
ggaggaaaaa gataaatatt ttcccttaatt tgcaagtata actctttatt tttttttctt	2340
ctaaaaagat tgaatttcta tttcaattct atggtatagt atatgtgaac ttattgtcac	2400
tgatactggt ggttatttat tgtaaacatcc aaacataag tcaaattttg ttaactgaca	2460
gttgctctac aatgggtgtct ctatttttga attagatgtt gtgttcattt tcatgaagat	2520
gttataatat actagtaata acttaaaaca acaagtgtat agcttatttc ttatgcctat	2580
gtaagcattg tgttctgtcc ttttttcccc ctccaaattg atatgtcct agtattgacc	2640
actccgcatc taggacttaa agagaaatc ttgcgcttag actttcatcc atactcctca	2700
attccatggg agatcatagc tgtaaatgtt atcttctgat ggtctgatac caaaaagcat	2760
ctagtttact ttgtccttgc tgcaattagt atgataagct gcctgaaact aattgtgttt	2820
gattcagttc tcattttgaa gaaataatgc actgactagc agtctagcac tattgttata	2880
taggattcta ggtgtattca gtatgatagc ccgctatagt atgaagactt aaaaggctag	2940
gctgctgggc actgacatct attttgggta totcatgtca aaatcaatta atacttagga	3000
tgcatattga agaacttact tatttacaac ttattttgaa ttacctcatg atataagctc	3060
ttgtgtaaat gtttctgaaa gcttataaaa ttagtttatg atcagttcat aagttatttt	3120
tagctcatta atttcagtaa gctctggata gcttatgaaa atagcttaca acttatatta	3180
aaacagttta acattatttt ctcttcactt atagaaataa ctgttacata agtgcattga	3240
tgataaacac tcgattagta agtgcttaat taagattttt aaccaaaca gcccttatgt	3300
agtggttttg ggtgtgttg ctgcaatcag tataataagc tggcaagtac aattgatatt	3360
tgtttgtatt cagtgttata tgtcttgtgc ctttgagtca attaatagga tcacactgtg	3420
tatttagtta acctgataag taaagactga ttactgaatt aattgcattt gcagcattgc	3480
gagcccaaac tgatatttgc ctcatgaaat atgaagatg ttgattacta atattccact	3540
ataacatata attgcttgtg ataacaaatg gataatcact ttttggtttt ggttttaca	3600
atagaaagct tgcattgtat catgtttgtt gacgatcagg acctactaat gggaatttag	3660
agaagactag ggtgtgttg ccttttaact tgtgctatat tgtctttct catgctatta	3720
cttctcaaca ctgggtagaa ttccctccaat ttgctttttt atctcttata aagtaaatgg	3780
cgattaggaa gcaattaata tcacatgata ccggagaaga atccatttta ataattgacc	3840

ttggtaattt ttgtgaatgg gagtgcgtaa ctcaaatatc agataaatga tgaaaacata	3900
acatcatgaa aatagtcagtg gttacaaatg tttatgttaa cggttcattt atagatccaa	3960
aaagggaaact aatataattc aatgattga taaactaaag gacaaatttt acatatgggt	4020
ttttaggat ttgttttcac tcagaattaa agtttcaatt tgtaaatatg cataatcatt	4080
catggcaggg gagaagtttt tttctatgt ttagaaatca tgaatcctaa gaatcatttt	4140
tttttggaag cgagtaaaat ctgtttgggt gaccatagat aaattttcta gaaattaaag	4200
agttatataa tatttttatt aattatttat cacatcaata aaagtaatat gatattttta	4260
ttataataaa agaaaaatta aactgaaaa agtaagtta ctacaatgag aaaaattttc	4320
ttgttcgtga aaaaacttgt taaaaaacat ttcctaaaaa tattgttctc taggaatcct	4380
atttttaacg gatatacaaac atgggaaact aacattttta tctcaaaagt ttcataaaac	4440
tcaatttttc ccaccaaacc acctcctcaa ttgtagacct gtatttcaaa tattaactat	4500
atgaaacttc aatttctaatt gaaaaatac aaacaatac ttatgaatta tatgtaagtg	4560
ttgtcctaaa ctaaactaaa ccaatcaaca aagaaaataa ctcaaatctt aattgaaca	4620
aatgaaaaga tcatttcact gttctagata aattttatga tatataatgt aaagtgaat	4680
taatacctaa attaagatag tacacaggca aacttcaaga aaagagaatt aatgaaagat	4740
ttctacaatt acaattattt aattcattga atagaagaag tcccaaagcc ttttgttttc	4800
cagtgaaga accaacaaaa aattggttga atagagagca ttgcgcgctt tagttgattt	4860
gatgataatt gaagataaag atttatttat attgtttcgc tatatatctc tatgtgtttt	4920
tgaggcgaag tatgatgctt gccgacctca cttgctaatt agtggatcc agacaactcc	4980
caaaccttaa aataggatta aagtttcacc ttttgtttga aatatttctt acagaagta	5040
aaagcttcca caattcacgt taagaatcca attctgtctt gtatttgctc catctcacgc	5100
ggtttcgtaa atcattagaa cccacttct ctcttttttt ttttttttt tctttctctt	5160
ccactttctc tctctgtcca aaccatagag ttttactcac acctcttcag ctttcaatca	5220
tggcaaatgt aaggcttttt atbaatttct tgctttctat ttgcaactat gagactgaga	5280
aggcttttct tgttgcatgc catcatcagg ttgtggaagt gaaagtcggt ttgcactgtg	5340
atgattgtat caagaaaac ctcaaggcca tcaagaaaat tgaaggttct taattaattt	5400
tctcaacacc ataacaacat aatcactttt tgtgttaatt agtttaaaat ttgtgttctt	5460
tgcagatatt gaaacttata acgtggacac aaagctgaac aaggtcattg ttacgggaaa	5520
cgtgacaaca gatcaagtga tcaaagtcct tcaaaaaatt ggcaagaatg caactgccgg	5580

ggaagatact caaaccaata agtgaacttg tttaggatag gacccttttt ttcctcga	5640
acatactcctt ctaaactatg tactctgaaa ttcaataatc tataatattg gaatagttta	5700
atthttattct tgaatttcaa ttgcaacact tgatatatat gttttttctt ctttatttaa	5760
ttttaaaaga gagatttagg aaccattgga tgattgattg attatgcagt tctttagctg	5820
ttctgaagtt tcatttcact aacctatgta tccttacatt taatgtaact tttattatgt	5880
aattttgttg acagatttta taatgtaaat tactgacacg aaatttctat ttggattaaa	5940
aacaacacta aactggataa aagtcttgtc ttttatttgg gaaatttaat ttatatatac	6000
cgataattata aagtttttct acactgatat ttacttgtaa attatttttt ataataaaat	6060
tgttactctt taaaagttat acctaccaac taatttttaa ttaattgata atataaaaaa	6120
ctaacagc acataaaacta aaatctatct aatgtttataa tcgatttatt atttatataa	6180
gtgctgataa agttgaaaga agcaaaaaag tgaacactaa caacaccaag taaaatatta	6240
attgttggct acctctttct ctacgtata atagcgtgaa gtgttactac tgaacctggt	6300
tgctgatag caaaaaaatt atttttttgt gacggtctta tccttgagtc gtcctgttaa	6360
tctagctaaa gttaaatttg cgtgcgttcg ctattatata tatataattg acaatattaa	6420
ccattaacca caatgcattt gtatggcaat agcttttctga ttacgtagt tctgtggatt	6480
tgctaaaagg tctcaagtta atatgatttg tctctccac acccaagttg tataaatata	6540
gataaaata tttcttggtt ttgtacaag tctatagcca ccttacttga aggcacaaccg	6600
ttaaatatgt caaatgcaa atacaagtc aacctcttat tctcttggtg atacttcttc	6660
atgggtttag agcactgtta tccttctctt tgtttttctc aaggaaatta aagtactttt	6720
taacaataag cacaagaaca actaaggttt ttttttctt ccttgatatt gattcacaaa	6780
tatggcgcgt gtatctgaaa ttggaacaa gacatttgta tgccaaaatc cccacagttt	6840
tggtcacttt gatgtctggc accgaggcaa cccattagaa tccccactt gtcttctctt	6900
tctgcaggct tctatgatga ccatagtcac acagataatg gacgcatgtc tcaagccact	6960
aggccaatcc tccttggtat cccagattct tgtaagcttt ctctacattt tacaagggtg	7020
ttaatgatgg tcactagaat cagtgtctta tgacctgta attctaaata tctacagaaa	7080
tattagtaac actcctttga atacattttt tagtatagat taaaatttgt taaatactat	7140
aaagtttggt ggttatgcat tttacaagaa agacacttgt agattgacag gactatgata	7200
gctgtaataa tttctctgta tagtgtgtgt tgctaccact cctatcacca ataattgtaa	7260
aatgtgattt gtgtatcaat aatgtaaaaa tagcgtgttg ttaacaaatg atgcattgta	7320



acaggggtgga	gtgctgtttg	gcccccaat	gttggggaac	aaaaacattc	tggggcaaac	7380
cttatttcct	gtgaagggtg	ctgtagtgct	tgaacagtt	gcacgtttg	gccttatgtt	7440
cttcttcttc	atatgggtgtg	ttaaaatgga	tgttgccaca	ttgatgaaga	ctgaaaagct	7500
ggccatcact	gttggttatt	ctgtgttcgc	atttacattg	gtaatcccta	ctggactagc	7560
aattctgttg	aggaatatg	ctacaatgga	cagcagcctt	gcacaggcac	tgccctttat	7620
ggctctctca	caaaccttaa	ctgtattcat	tagcatagca	gtgctcttga	aagatctcaa	7680
agtccttaac	actgatattg	gacgtttaac	catgtcagca	gcaatgtttg	ctgatatagc	7740
tggctttacc	ttgacagtga	tcatttttgc	tgtactgcag	aatcagagtg	gtagctttct	7800
gacactagca	ggcctccttc	tgtccgtagt	tgcactcttt	cttctgttta	tatttgtgat	7860
gaggccagca	atactttgga	cggtaaagta	ctcaggtggc	ggttcgggta	atgaaagctg	7920
cggtgtttgc	atctttctct	tagttctctt	atctgcattt	atcagtgagt	taattggaca	7980
gcattttatt	atgggaccaa	taatttttag	cctggctgtt	ccagaagggc	caccataggt	8040
aacagctttg	ttgagtaaat	tggagacaat	ctgtatggga	tttctttatc	caatctacct	8100
cgctgttaat	ggattgcaaa	ctgatatctt	caaaattgac	ttgcaatctc	tgtggattgt	8160
gggtctcata	ctgatgtgtg	ctttcgttgt	aaagatttgt	gctgtcatgt	tgccaggata	8220
cttttacaat	ctacccatga	aacaatgttg	tgtcattggc	ctccttctaa	atggaagggg	8280
tatagccgag	cttaccatgt	acaatatgtg	gattggaagc	aaggtatgta	tctatgtatg	8340
tgcataatg	attgagacta	gtgaatttaa	tttcaaatg	atgccaaaaa	tgtgtttcag	8400
gtcgtcaaaa	ttagctctaa	ggcactggga	ttcactgatt	ataaaaaata	gaataccgaa	8460
taccgaaatt	atcattttta	aagtatatgg	atgaaacaac	aagtcgttgg	ttagtgtagt	8520
attggacttc	attcccttaa	acaggatctt	agattcgagt	cttgtagatg	aaaaaaatat	8580
aattgggagg	gacaacccta	ctagaggtgg	tcagccagat	tctcaacag	agattgatta	8640
tcagcaaagt	taatgatatt	ccgtacccaa	taatatgggt	aacaaaaaag	tatatggatt	8700
aagactaagg	ttaattgagg	gacataataa	aacacatttt	acccggttga	ttaactaaca	8760
gtctaatact	tgttctcttg	tctatggtag	ttaatatcgg	aacaagagtt	cgctttgatg	8820
gtggcctcca	ttgtagtagt	aaacgctatc	ctagccacca	tcgtaaaaata	cacatatgat	8880
ccttcogaac	aatatcagac	cgggaagaaga	tgacacattc	agcatactgtg	gcgagatatg	8940
gagctccgag	tcattggtgtg	cattcacaac	aatgaaaacc	teccaacaat	cctgaacctc	9000
ttagaagcat	cctatgcaag	cagagagagc	aagattgggg	tcacagcatt	agtcctagtg	9060

gagcttcaag gaagagccag gcctattcctt gttgataacc aaaaccaact ccatgatgag	9120
ctgcgctcaa tgtcttgc aa cgcaagtcac attgagaatg cattgaggca atatggacaa	9180
cagaacgaag gatattgtatc tgttcaatct ttcacttcaa tctccacctt tgaaccatg	9240
tatgatgata tttgtagaat ttcatttagag agtgggatcca acattttgat cttgccattc	9300
cacaagaggt gggaaattga tggcaccgtt gagattttctc ataggaccat ccaaaccatg	9360
aacattaatg tctctcaaa gggcccatgt cgggtgggaa ttctagtga taggagcatc	9420
ttgaatcctt ctcttctact cttgatggct agagcagcat tctatgtctg ggtgttcttc	9480
attggcggtc aagatgacat ggagacatta gcctatgcc ctagaatggc taggcatgaa	9540
tggtgtgatg taacctgtgt aagggttctt ctatttggag aggagaattc taaagataga	9600
aaacgtgaca gcgatcttat agatgagtat agatactata atgccagaaa tcgtagggtt	9660
gagattctgg aagaatttgt gaaagatggg atagagatgt caacatgtat aagaagattg	9720
atagattatt ttgatctagt gatgggtgga agggagcatc cagagagtgt tatttttcag	9780
ggacatgatg aatggagtga gtgtcaagag ctagggatca ttggagacat gctagcatca	9840
ccagattttg tgaccaaggc atcgtttgtg gtggtgcaac aacagagaat aagagggagg	9900
cttgttacac ataattgtga tgccactcca gtgcctaatc aaagagatca gctctacat	9960
gatgttccaa ttcatgaaac ctttagtctt tcatgtacta tttagtgga caaatatgac	10020
aaaatgtaga tagattcctg tgtttagtga attttccttg tcgatataata tatcatgcaa	10080
atcacagatt ttcattttaa accacaatat aatgttactt tcgacaaagc catgaaaaaa	10140
taaaattgat catcacatat ggttagtact tgatactaat acattgccct cacttaataa	10200
ttttgaatta aaaaaaatac ttgaatagtg gaatgtgact aacttcctta aagggtgtga	10260
gaggcataaa tggaaaaatg tgattttatt taatataaat ggctaaatga gtaaaccatg	10320
tttggttctg cacaatagtt atgaacgttc aaagtctgtc ctaccaact ttgcattgata	10380
ttcaacctta cgggcattga atgaggaaat cccttaatt ttactaaaa aaaaatcaa	10440
acaaggataa aataaaaaa cagataaaaa atccttaagc tacaagtcta atcagaaga	10500
aaaaaatatg atccagaatc atcaaatgtt tacaattcca attaatctct ttgtaacacg	10560
taatgttaat tcttttcgta aattaaaaaa attcaactac atgttgtgta aatttacaaa	10620
aattatatac acataacaat ctcaaatcaa aaaataattt ctgaatgctc atccacagaa	10680
aggtttaaac tgcacgtctt cacaagggtg attccttaat acagtcagtg tatatatata	10740
aacctaaaag taatttatta gaagggtgcta ataagttggt gaagcaagcc aatctttcca	10800

taaagcaatg gtcattggaag gtgatacttg gaaagaaagt acctgaattt cttggtaaag	10860
caagaaatta tgtaaacaaa gacattggcg tcaggaaaca gcattctccat ttttaagtgc	10920
aatttatggt gcgtattcaa acatcatgtc aaataaaaga aagaaattct tacgactttg	10980
taccatttca tgcgttaatt gaggggtaca cttttttttt ccaacaactt gaggggtaca	11040
cggtgaacag aacctaaacc gttctcgtcg aataataccg attcgacaaa taaaaaatga	11100
ataaattata ttggcaaaaa aaaaaataga ataaattata ctttattttc caactatttc	11160
ttactttttt agttttctct ctctctctat aagtatatata ttatatata aaaagacgaa	11220
attcgtaagg caatcttatt ggtattataa ttttctctac tgattatgtc taaccattta	11280
tacacacaca cacacacaca tatatatata tatatatata tatatatatt attacttggt	11340
aaataaaatc agaaaaatgt tgtaatactt tcaaaaactg tagttaataa accttaacta	11400
aatcaagcaa aaacaatgga taagatggaa gtttagtgat acaaaaatat atacaggat	11460
agtgagaata aaaaagttga ggaagtgtga aatctacgtg aagatgaagg atgaaaattg	11520
gttaagttag gttaggttag attttataaa atatgtatga gttgatattg ttttttaagg	11580
ttcgagctta gtttatatat ttagttgata tagactacat ttaaaagact gacttaaaag	11640
tctctttaa atacataata ataacgaaa tggattaagt tagatttggt ttttttctt	11700
tttactgggt agttagatta ggttaactct tataaaacat gaatttgatt tgtttttttt	11760
ttcaaaaaa attaaggttc aagcttaatt tatttagttg atatagacca ctttcaaaaa	11820
tctgacttac aagtctcttt agaattcata atagtacac ttgattaagt tagattagac	11880
tttataaaac acgagtttga tttttttttt taataataat taaggttcta gcttatatat	11940
attatatagt tgatatagac tactttcaca agtctgactt aaaagtctct ttagtatata	12000
taataatata acottttaab ttagttttaa aatttgctcc taaataaatt aataaatcca	12060
aacttatata caagttaata ggtttaagtc ttaatatata tatatatata tatatatata	12120
tatatatata tatatatata tatatatata tatatatata tatatatata	12180
gagagagaga gagagagaga gggtcatttt atacagatga gaaaatttaa atattattat	12240
gaattgtcaa aattaaaata cacacatgcc acatgatattt cttaaaaaaa ttacgttaact	12300
tttttttaca aaaataatca tatgggttta aaaactaatt taaataattt atataaact	12360
atatacagtt aatttgtttc ataaaaaag tatatcagtt attttacaaa attataagta	12420
ttcataaaat aaatacaaaa tgataagtc caagtgtatg gatcagctta tgcgatgttg	12480
ttccaatgta actaataatc ttaaattcga gtattgaatc gaatatgcaa ttcattttaa	12540

tactttaaga gataatttgt ttactcgtaa taattttatt tgacttaagt aaagttttct	12600
catataaaaa atacatatag tctataaaaa aaacattttc tgactaaata tattttcaca	12660
ggctccacac aaaaaggaaa taacaaagtt ttaagagaat gtatatattc actcatcagt	12720
ttgttaaagt taaatataa atggagtaga aattgagaga aaagagagga aatatataa	12780
ataaagggtt atttgtataa ataaaatgtg aaggaaagaa ataaaaaact ggtgagtatt	12840
actcaaatt atctttatac tgatgaaag aaaactgata aagaagtaa acactaaaaa	12900
aagagaccaa cttacaacta acatttactt cgagcatagt taattaagtg ttggggacat	12960
aatattgttt ctataactaa gtcttacgta gattaaggtt tacgagacct atcctaataa	13020
catgaattga aactagtatc acctgtcttc tgggctcaat cctggggatt cataaagaca	13080
tttcttgaac aatcaaagg gttatatgaa tgggttaacg atcattaatg tgatatgtat	13140
tgacaaccga tcaatgctag atatatagtc ttaaaatcct gtatcagctt acagacgact	13200
aatatgatgt aaaattcctt agtttttaagt gcttttgaac atcaagagac ttaaagttcc	13260
gtgttggttg acaacaatat ggtatgatat ataactcttc aacgtaaaca cgaaaaaaa	13320
acttataatc tcgtgccaat catcgatata gtacaaataa taaattaaaa tgcaattttt	13380
ttcttgttct tattttttct tattttctct aaactagata ctatcgaaac cattctattt	13440
cttatctgtt tccattatct tactttctac ttattttcat tactttatct ctttctttta	13500
tgttttatct cactttattt atcacctatt tctttcttct ttaccgaata ctaacaagc	13560
cttgtgatcc gaaagccga aacaatcatt ttttatgaaa cagcttacac tctggtggtg	13620
tgttgtgtat agttaataa gcttttaaaa atatggtaaa ttataagggt agggaccaaa	13680
atgtgagatt gaaaataacc gtttcatata ttattcaaat aaataaatgg ctaaaattga	13740
atcattctcc gcataataatg accaccattt tattttatta atatattctaa caattatttt	13800
taactccata taggcatttt tgacctcat cttaaaactc acctcaagaa atatattagt	13860
attttaatta aattagtact caacttcaaa ttaattatta gacaagtgtt gttttttaac	13920
catttatcaa attaggaact ttatgtcacg ttatcctaaa atcgttacat aaatttttaa	13980
tgtcacgata caatttttta gaagaaaaat ttgtctgaaa cccatatgac atgggatgca	14040
ttagtcaag taacacttcc taaatcatca acttagttag tggcatgcaa catggcggtta	14100
acctattttt ttttttctag aaaaaaaa acatataaat atcaccagct gatgtcgcgc	14160
caccttaac gccacgcca gtgtaggcgc actatgaaat caatgcagtc agttttgtca	14220
gttcagactt gcaaaaaaat ctccaaaatt atatcacact taataaattt tataaagtat	14280

tatgcataat	atTTTTTTaa	agttattatg	catagtatct	taaaagtttc	attaccacta	14340
tttttcatta	gataaaccta	tttaataact	tgtgtaggct	catatataac	caaaaagtaa	14400
tcttgaatgc	tagattttaa	tataataaaa	tttattttat	cttattttat	tttaataaatt	14460
ttttttattca	tcattgaaaga	tattgtaaaa	ataattttat	tttccctttg	agattaaaaa	14520
aaattaaatg	atatgaacca	accttttaat	taatatgaaa	ttagttattt	ttattttatat	14580
ataacttttag	ataacatatt	tttactcaaa	agactaaaat	'aggtgggaat	gatggtcaga	14640
ttaagtgttg	tggaattatt	gaagtcttat	attcaaaagtc	ctaacttata	actatatttaa	14700
atacttataaa	gtttttttatc	atctataaac	atcatattca	attctaaaaa	aattgtttct	14760
actaaaaaaa	aatccgtgac	tcagaaaaaa	aacattactt	gggttgattt	ttgtagcatc	14820
taataataatt	aggtgttgac	tctaattttt	atittgacgg	aagaaagttc	agaatgacga	14880
gaaagtgttg	aagcaatatt	ggtgttcgac	ttgtcgtgtc	tgagacgtgt	acgtgtacct	14940
tccttaccta	aaaaatgaca	attaaaaagt	gtttattcgg	gacgcgccta	accaacttca	15000
caatttttta	agtaccactc	cgggtgctat	gttacatggg	cgtcagttct	tttttttttt	15060
ctttttaact	aagaaaaaatt	agtattttta	ccaattatat	aattgtgaga	ttaaatattt	15120
aaaatatata	attaaaaattt	ataataaata	ttttttttaa	atatattaat	tatatttgta	15180
tttataatga	ataattttata	taataatata	aaattatttt	tagttattga	tattaaaaat	15240
acagatgaaa	aaaatatatt	gaaagagatg	aaaggttgag	aattttttta	aaaatattgt	15300
tattgaaaaa	ttattagaaa	aattcattga	agaatagttt	taaaaaaatt	ctcattcaag	15360
ataattattc	ttatatatat	aaattttata	atgatttcat	gtaatacgaa	aaattgttaa	15420
tcacatgac	aatcttagcg	aagtggcaaa	actgccttgg	taaaaaaatt	tcataagata	15480
atccctttta	aataattaat	ttgtaattaa	tattgtgccaa	tcctcgttac	cttcgtttca	15540
tttagcttta	agaacaaatt	tatgtttcgc	atcacgtaac	gtgtgtgttt	ggaagcgttg	15600
aaaccacatt	taatgaaaga	aaaaaaaaac	atattctaag	acacaagcaa	caaaaggaag	15660
gtttcttttaa	cgtcagttag	gttgagataa	atgcgtacgc	aaaacagtt	aaatgattag	15720
cttttagaaa	aagtttagag	ttataagaaa	acatagggaa	atgtcactaa	tttgtttatac	15780
tttcaataat	caagatgtta	gttttctctt	tcaaaaacat	attaacttgt	tctctaatta	15840
ccattttttaa	tttttaacta	cattttattat	aaatacttta	gtgaaaaact	catagcagct	15900
attccagttc	tgttcatgta	aatatcgtag	aagataattg	catttttttc	cttttctttt	15960
tctaaaacaa	gaaacgtgtg	aggaatctta	agaattaaga	tgctaattta	aaaagttgct	16020

gagttagagc ataaaagttc aaataaaaaa taaatgaata gacaaactat taaattatta 16080  
 atagccttag ccttgaatct gatgcagacg tgtatggcaa tggacataga agcattaaat 16140  
 aggcctcgtt acattcaagt ttcaaccaa ttggcaggga aaatcctctg atactgttgt 16200  
 tttcctgaaa ccatcacaat ttgtttctca atcatgtcaa cctcatcctc ttcccaaaagc 16260  
 ctcaaaaattg gcatagttag attcggcaac ttggccagtg ttctggccaa gacaatgata 16320  
 aaacaaggcc acactctcac agcaactctc cgatctgatt actctgaact ttgtctccaa 16380  
 atgggcaccc attttttcag gtaagtcaaa ccaaaccaa coaatgcataa atacatacac 16440  
 acttgaccca ttttgcctga aatcccacgt ggatcagtga tatagtcaaa atagtgtata 16500  
 taaatagaga acaatttttc acttacgagc tgattttgtg aagttaaagt ctaaaggcaa 16560  
 attctaagac atttcatgtt ccgtatgtca aacattacgc agggatgtca gcgcattcct 16620  
 taccgcagac atagatgtca tagtgttgtg cacatcgata ttatcgctat ccgaggttgt 16680  
 cgggtcaatg ccaactcacct ccctgaagcg accaacgctc ttgttgtatg ttctttctgt 16740  
 caaagagcac ccaagagagc ttctactcgc agagttgcca gaggattcgg acatactctg 16800  
 cacgcaccca atgtttggtc ctacagactgc caagaatgga tggacagatc acactttcat 16860  
 gtatgacaaa gttcggataa gagacgaagt tatotgctct aatttcaccc aaatttttgc 16920  
 tactgaggtg ggtaaatatc ctttgtcaat acccatcaat cagcaagaaa gaaagaatca 16980  
 tttttttttt tttttatttg gccagtttaa ttatgttaat caagaagaaa gaaacagaga 17040  
 ggggtggaagc taagtaactt cagacgtttg catttgataa atcaagatac aagataaatc 17100  
 tatgttgtaa aaaatgtaaa gtctcagtc cactctaac agaggtaagg attaccatca 17160  
 ccttaccctt ataagttatt tgttgatttg agttaggcct aaattttaaac ttcaaaagta 17220  
 tcaaaggtta tcttgatctt attaataggt cacttatcat attacctaca cacaaaaccc 17280  
 aataatgttg accgtgagag ggtgtattga gaaatatcaa cagtttgatt ttgtgagggt 17340  
 aagctaggcc aagtccaaat tgaaagaatc taagcataca atcctacagc caaggttata 17400  
 gaacctcaac gtcattattg aaacaaatg catattaaac aattcaatct tcaattctcc 17460  
 cctggactga gtttgaacta attcttcata agtttggacc attgttgctt gggtttacaaa 17520  
 agaatacaca aaatcagttg tgaatgcctt acatttatga ttcatcggg gtacatctat 17580  
 ttcaactgtgt ttgttactt ttgtattacg cagggttgca agatgtgata gatgtcctgt 17640  
 gaggaacatg acagagcagc tgctaagagc caatttatca ctcacacaa ttgcagggtat 17700  
 gcagcttcct acatatctaa taaaccattg agaagcacta atataaatgc tcaactagatg 17760

taacttttgg ctgcttcttg atcaggacac tgggagaaat ggatattcaa tccacaccta	17820
ttgacactaa gggcttcgag acacttggtta aattggtaaa gagttgttaa catttcccct	17880
actttctcta aaaaatttcc tttatatgtg gttatgtatc atatggaaaa gttacttgca	17940
gaaggagacg atgatgagaa atagttttga ttgtatagt ggattattcg tgtataacag	18000
attgccaga caagaggtaa tggaaactgcc aacgaaatgc ttatttactt ttaaattcca	18060
ttttaagca ctgaaccaag taccocaaga tagactcatg aattgtgaaa aatatgcagc	18120
tggaaaacct tgaacatgcc ttgcacaaag tcaaagaac gctgatgata caaaggacga	18180
atggggagca gggctcataa agaactgaaa gttgatgcac atttatttta caagatattt	18240
tctctaacte tcaaatatcc tctgacagtt ccaattataa attactctta tttcagtttc	18300
cttttaccaa aattgaagtt caattaataa accaaagaga ctggtatatg ttcaatcaca	18360
tgcgataaaa atgttccacg tcttctgtcc gagcagattc tttgtaattt cataaagtta	18420
gagaaaagaa aaaaaaaca gacatttagt cgccaatgcc taaaaccata taataactcc	18480
acagtttggg tctctgaatg aattcccttc attttaatcc aaatctcaac tacctcttca	18540
attctaaca aataaattag aacactacca agtgatcctc tggggctctg acatgagcct	18600
tctagtctta gcttttcaat caatgtatct acagagcatg tattccattt tggattatag	18660
agaaataatg taaaactttt aaccaaattg tactgcaaat ctaagaagag ttcatgtctc	18720
catgataaat tgataatact acataagatg tacaactgct gattttatat atcattttaa	18780
caagacttgc caagagatat atcccttaaa gccaaagaca cttatgttcc gatttgagac	18840
acctctattt attccactta catttgaaaa ataaaaatat tacattcacc aaactgggaa	18900
atgggaataa tcaaaacgta tagaagttag ccgtggaagg aattgtaaac aagttatgac	18960
aaaaccaga acctattcct ttgtgcctat ttaccagct ttcaagaat agtacaaaat	19020
taccaagaaa aaaaaatgat gcaacgtatt ttcaccacat ttatttttct ctctcatcaa	19080
tgctggcttc tctctgttgc ttgcataaaa ttgcagcagg attcggagga tctgatgctg	19140
atggctctgt ctttctaaat tgtaggccag taaagcgcg taggtcaatc cattgctgca	19200
gcaactcacg aaaaacaaaa attttaaaac tcctagtcaa atttagaaat gtatgtgcag	19260
atgaataatt cacataaaac taaattatac tccattgggt ttaaaacaa tgctactctt	19320
gattcttgag gtttttagct gtttcaaatt atgtcacttt agaacaatgaa gacgacattg	19380
cttttttctt tgtcttttat tgtttctcta actaaaaatg aaagcaataa tggcggaaga	19440
ggataatata gtcagacgaa cttagggttt ccttgaaatt aagaatattt aatgacttta	19500

ttaatcatag tagaaaaacat taaacaatta ttgtgaaatg gaaggggaagt atattgcaat	19560
atggcaaaag caaatgctaa aacagtgtct ttacaatatt ggctaaaaag taaaaagtca	19620
aaagaaaaaa gttattagaa aatgttacta ataatacatt tccattgtga ttttcaatac	19680
atccctaagt taattttctaa taaaagaaat tctattatta attcctcaag accttattac	19740
aatgctaat gagtttccct aagactcttg tcaagggatt aaaaaataa aatatattaa	19800
atgaaaagta atgtattaca tgtgttactt ttcataaaat gctctatatt ttgggcttaa	19860
ttatactttt gatcttctta ctttttcaat ttgttaaact ttatccctct atttttttct	19920
cacaattttg atctctaatt attttaatta tccattaact taacatccaa tattcgataa	19980
atgtgtctgac atggcagtgat agaagagtgat catgtcaaca tgaacgtgtt gacatactag	20040
caacacgtta acaagacctt tttttattcc ttaacaagca ccttaattgc actcgtagc	20100
aacaccttg taatattaag atgacacatt ataagacaag cggaagtta tttcttccct	20160
taacttgtga tttattgttc ctgacaaaa gcacaagctg acgcaattat aactcctcaa	20220
aacacatatt tccatattaa cacttgaatg tgaattccac cactttaaaa agaaggaaaa	20280
attaaattag atttttgaga agaattatag tgttcaacca taaataaatg aaatccactt	20340
actgacgaga ttgtgaacat tgagcattga aagtgaacaa aagcagatag aagataagga	20400
aaaaaaactg acctgggttc cccaatatga attgactgtc cgtgctacaa aagaaagcat	20460
atttacaag aatgtttgag aagcaaaact ttcagaaatc cgatgctcca ctatcccgac	20520
aattatctgc tcaacaaaaa tattataatt tccgtattaa tacaaaaaaa cagttccagg	20580
cattaaatgg atgtatattt gttgaatatg atataccaa agatgtagtc acagaacaat	20640
caaaaaatcaa ttttaagaaa gaaagagcct taaatagtta ttcagcaagg tgcagatgaa	20700
gaagtagaaa ggaggagaaa attcacaatc cagatataat agatattatt gaaccgagta	20760
atttttttct cagcattact tttaatagca tccaggggtt actaacataa tatbtgggtg	20820
gaaaataatg gagaggaaa gagggcaagg atttggaagg aaaggagaaa tggaggggag	20880
taaaacacct cctctataca ttttggctcc cttccaaaat tgggagaatt tggagaggag	20940
aaagttttac atgaattgga ctaactatc cttaattggt ttatctatt atgaggatat	21000
aataaataat aaattatttt attttccctc ttatttccct gtgaacaaa ccaagtgttc	21060
ctccctctca ctccctttct ttgaacaaa tagatagtat acgatgtagt ctaatcttct	21120
ctattatgct tgcctaccac catctacttc aaactctgcc taagtggccc gaaaaacata	21180
gtgtacaaat tagttattca agacacagct ggtcaagtc tgctttcaaa atgtctacat	21240



gaatttccaa attgtatttc tgtcctgctt caatactata tattggaaaa aaatcaaggc 21300  
 atcacataat caactattga gtattgaaac ataaaccatc gacgtgacaa tatagaagcc 21360  
 ttctattcct acagaaacta ggttcaatct tccatgaagt ctattgctac atattagtta 21420  
 aaatgaaaac tttatgcaac aaacaagttc tctatatacc tgataacgga gatttgcaga 21480  
 tattccaaga aagcatgcat atataaatgc agttttgagt attggggacc tcataaatccg 21540  
 ttgttcactg acaactgctg gattaaggac tttacgaatt gcataatagag aatttgatga 21600  
 agccacagct cctatgcttg aaataaatcc aacactagca agttttaaac caccaaatac 21660  
 aactgatgca atcctatgat tgagattcca gtttatccct gcaggattct tttgaaatgc 21720  
 attgtccggg atggagccaa gaagtcctat tagagaacca atgttgtcgg gtgctttcat 21780  
 ctcatcagca tatgaaagga atgacaaagt tgggtgcagga agccacactg taaagaaatc 21840  
 aacaactgat cctctgacag tgtccgtaat aacatagta atctcttgga aaaaatttcc 21900  
 ttccgctttt tcatactggg ccaataatgt agttgttatt gatatagctt cttctatggc 21960  
 taatctgtgc aagaatttgg gatctgcca caatctttcc ctgaatccct gtataaaaag 22020  
 aaccagatag tgaagggaac aaggtgatga atgattaaaa gcagaatgaa ggtttcattt 22080  
 tcaataaaaa aattacctgg aaacggtgag tgagctctga aattagagga tactgtccca 22140  
 gatcaaagaa gttctgcaat acctctgggt atactaaacc aagatcaatt cccttttgaa 22200  
 gatcctgaaa ttgaaaggaa aatatttctt atgttcacca ccaatatgg cacacactat 22260  
 cctgtctaag gtatgcacat aaaagagtaa acagaagcag agaaaaaaga actagaaagt 22320  
 agaaacacac taactaccta atcatcaaat gcaagtatgc aactaatgta tgccaatta 22380  
 tagaccagag gtactttatt ttaagagaaa gaaaacaggg tactttattt taagagaaag 22440  
 aaaaacagaaa tagcgacctt attatcacia ttcacccatt ggaaaagttt ttatgtctct 22500  
 aaattatata ttgctgtgca atgaatcttt cctcaaaagg aatatgaatt taaaggaaaa 22560  
 gaaagatagc acaaagacag cactacaaag ttgcaagcat tcaattaaaa tccccacac 22620  
 cagtaggttg agctgcata tttgtgtcaa ttaataaat gcaaacaga gatatcaatt 22680  
 aaaggataaa ggacccattt atttaagctt ttaaaaaaat attttttttt acatatttta 22740  
 tgtaaagtta ttttatttgg ttacaataat taaaaaatgt actttatatt ataaaaagta 22800  
 gttataattt tgactttttt tcagctgcta ctcaaagtag cttctgaaaa taatcatata 22860  
 gatagataga ttctgatttt ttttctaaaa aaaaacttaa acaaacacac taagaaattt 22920  
 tagaagtgat ttttcatgaa aaagttgaa acaaatgggc tctaaaatgc tcctgaaatg 22980

ccaaagttaa	ttgcatacaa	aaaaataat	caataggtag	tggcacaaga	cacctagtaa	23040
tatgcgaaat	ctcttatgtt	tgtatcacca	aaatggacaa	tgagaggaca	taacaacaac	23100
aacaccacca	aaaccttacc	ccactaggaa	tgagaggaca	taaagggcta	aaaattggaa	23160
ggagggtcta	cggggcaaga	agattaacag	tcaaacaagt	tagtaactgt	aattgttggt	23220
ttacctgtgg	gagggcatct	cgctccgccc	cagcagcatt	cataaccgga	gcaatctcag	23280
cacgggtcaa	gcaatttcta	ctacaggggc	tcgcagcaga	ataccacaaa	aaatcagcaa	23340
caggaacttc	tccttctctg	cgaatgaatt	gtctttcagg	gtcaagtaata	ataactgcat	23400
ggtttttctt	ttgtattttt	cctgaaattc	ttgctggcac	tccagttcct	ctagatccat	23460
atgtaacatg	gcttgaccca	gtgacaacta	ttaacatacc	agtgaccctc	ccatcaagca	23520
cattttgtaa	gataatctgg	gacatagaat	actcatcaac	tactcgagcc	tgtgcagaaa	23580
ggtagtagct	tggacacaaa	ggaatagaca	gatttttagt	actatcaaca	gaagatctgc	23640
gtgagataga	agtaaaagcca	gatatgaagc	ctgaaccagg	tggagggtgca	tatagtttac	23700
gttcacacct	tgaagcccca	cgaattcctt	ctgcttgagc	agttcttaag	atctgttgat	23760
gagaaaagtt	caagtcttaa	ttgtcatctt	tatgtgtctt	gacctaacaa	tacaagaact	23820
gagcctaata	ctacaagggt	gagttggcta	gatggatcaa	atgaaacctt	ttagctctat	23880
caaaaaccaa	aattgcaata	aagtttctca	ggttgcccct	taccttttca	acagattata	23940
ttcctttctc	atatcttgac	atgatcaaac	tacttgaggc	aacttttcat	aatcatatcc	24000
tcaatttttt	tatgtttaac	tttcagttaa	atatgggaag	gaatctttga	ggattatgtg	24060
tttattttcca	attcagggtc	tgttctttgt	agactagatt	tttctagtcc	taacgcaacc	24120
aaaatcctta	agggtcaact	tttacagcac	atagcttttg	cacttatatg	aagtcaattt	24180
gacaagcacc	aacaagtata	tactatacta	ctccctccat	tccaaaataa	ttgttgtcct	24240
aaattgtttt	acacagacca	agaaaaaaca	atagatagat	gaaagagagt	tgtagtttta	24300
caaagttaat	cttatatcat	cattgattca	tttatagatt	ttgtttccat	cattaatat	24360
ataaggaata	tacgtgaaaa	aatgtaatta	aatattatat	tcaaaactaa	ataaacaatt	24420
attttggaaat	aatatttttt	tcttatacga	caattataat	gggacagagg	gagtaacatt	24480
ttttctgtgc	tcctaattat	agccacacca	caaccataat	tttcagagac	aaaataaaca	24540
tttgaaagat	caacatgagt	ttggatgaaa	tttatgcata	ccttcagttg	tgtaccacaa	24600
gcaacaagat	gaattccatt	ttcgcgacag	tagctcagaa	taggttcata	ctcctgccat	24660
ctttgaggcg	gccaatgcaa	cgtgtaagac	ttcaagggtg	ctccatctat	cctgccatga	24720

aagtcaggtt tcaggacaag taatgcagaa ttatggaaaa gcaatctgac taagacaaaa 24780  
gagcttcaga gattaacaga aaatagttag ccagaaaaaa gattgcgaga cagaaattgg 24840  
tcgccacaaa aaagtgtgtc cttttataat ttttaattga aattttctta atttagctaa 24900  
catgactttc tacggccaca attgcgtttg cagacactta aaaaacttga tgttcgagca 24960  
aaaatcacgt tttatttatt attgatgtca attatttaac agttttatgt taggtttaat 25020  
aacagtaggt tgatgcaga.ggcctaacat taatcagaaa ttgaaaggca gtgttattac 25080  
ttcttatcca tatactgatt gagcggttcc tgaagattag cgggaaaaac ttcaagcgcc 25140  
agagacaata gtttttccct ctccaacacg cgccctatga aattcttcac aatctcaagc 25200  
tccaattccc tatcgtctcg aaccgggaact tgctctgctt cacctaaata caccactcga 25260  
gcattcatca acttctccca cactttccct ttctctttcc ctatcgccaa cggttctcct 25320  
atcacccgtc cgtcgtaaat cctcgaagtt atcaacttct cctcctcctt cttctctggc 25380  
ggctcctccg gcttcggcgc cgaaggcaga ggagattccg ctgccttctc ctccgccctc 25440  
gccgtcgcgc ccgagagcag gattgatgcg ccggcgacta agaaccggcg catcagcacg 25500  
ccgcgcgcgc tgctccgagc tcgagtatcg ccgtcccggt cggaaaccac gggattcgag 25560  
ggcgcggtga ccgactcgc gtggcagacg gaaaggctga ccggcgcgcg tttggcggtc 25620  
gagacgcgcg ggaattcgag gcctccgggg cgcgtatgag gaggggtccga tggcgcgcca 25680  
cgtgcgcgcg tggcgccgcg gaagtaaggc acgtgcggga gacgagtgac gaaagaagaa 25740  
gccggagtgt ggggcttcat tagtttcgtt ggcttcagtc tggccttatc atcaaccgca 25800  
acaaggttta agttttgtca gttcactttt ttcaactgcc acacaaaaca cgaacaagga 25860  
ttcctttatt ttcagcatta caagattcca taattattat tattattatt attttggaga 25920  
aacttttttt ttttctttc aaatatccct tttcatagac cattttaatt gacaaacaat 25980  
taaacattaa ataaaaaac atttattatc aaacagtaga tcaatttagg tgtgttagtc 26040  
taaagaaaaa caaataaata acaatgacat gttttgagag atagacaaat attcctctta 26100  
acataacata aaacatatat atatatatat atatatatat atatatatat 26160  
ttgtggtaga aagaaatatt agtttttagtt cgacatatata taagatatata gtttctttat 26220  
gtgaaattca tgttctttta tttattatgt gtttaattta aaagataaaa ttatataatc 26280  
taaaaataaa aaagaaattt ttgaattaaa ataaaaata agagttgtag gattcatatt 26340  
attttttctt caacacaaaa gacaccctaa aaaagttacc aatatacttt ttaacatggt 26400  
attttaaaag tatttttatt aacttaaat tattaataat tataattttt tttaatttca 26460

tctttttattc aatgaatttt cctataatt tgattgtttt caataaatta taatcattat 26520  
 taaaaaata tataataaag aaagtgatc aagaaaatat atcaagaggt tattcttatt 26580  
 ttgtcctctc ataaaatcct agacttaaaa tatcacactct actttttgtg ttcaataata 26640  
 aattcgtata ttctgtggtc gcaaaaaaat taaatgatat tgataattta catagtcacg 26700  
 caaactatat atagaaggcc tcaaatttag agctctacgt gaagcattgg tgtccataga 26760  
 gttacatggt gccaaattacc atacttttca ttattttgga attttcatgt aagataaaat 26820  
 cgatccgagt cagatgtgac tcaggtctga aacacaggtg cggatccaat atccatgtat 26880  
 gccaaataa cgaggcgtct actctaattt gattgaaaa aagggggcaa aaagtaaaat 26940  
 atatactgcc aaattccaat tcaattcaac tatcagacat tgtctaaaaa ttgaccacaa 27000  
 ttgaccaact gaagtgtac cttcttgttg atataaaaaa ctctgtgttg ctatgaatat 27060  
 ttttaaaaca aaataatatc tatgagataa aaatatcatt attgtcaaat aatgatgaaa 27120  
 tactatcatg tattttaaga agaaaaaata tataaatgaa gatgacttta ttgttaatta 27180  
 tgtgaattct attttattat ataaaaataa ttaaggttat gtttaaaaaa attagttgaa 27240  
 agttaaaaaa caaactaatt gataacccaa aacttttaag ttaatttatt aaattatasa 27300  
 tattttagat aattgttgtt gaagtagata aaaaatataa tatcacaaaa atagatatat 27360  
 ttatatgata ttatataaaa ctttaatggt ttatggacaa aagtataatt aggtattata 27420  
 attttatttt ttaattaatt ttaactctt gtaaattatt ttctattata tcttttgttt 27480  
 caattattag attttttcat gttgcatatt ctattattaa tcttgtaaaa tgtcttaata 27540  
 tttttaacca agtttgaat aaagttgaaa aaacatgcac ttaagtaaat attatacttt 27600  
 tattatatta attagtataa tagttaaaat aaattgtata atataaaaaa attcagaaaa 27660  
 taataaatta ttttaacatt ttttactgtc aatttcgtga agatgttgaa aaattacaa 27720  
 tggctaagc aaatagtata atataaaaaa attgtaagag gaatgagtga aaataaata 27780  
 aataataaaa ttatgttata tttaaaagga ataataagaa tatttaataa atattttaag 27840  
 aattaaaaaa taaaatataa aagcaaaaaa tttagaggcta aaactagag ttttaaaaaa 27900  
 gttactttta ataatgtttc agaaaaaat aaagttaca aaaaataact tattttattaa 27960  
 ataattaac aagtttttca actaataaaa aaataaaact aactaaaaa acgtatcaaa 28020  
 catagcctaa agcgaattta aaaaaaaaaa ttaaggtgga aaagcatcaa aattcaaagt 28080  
 tggatcaaa attaggatta actaaattta agcaataata tatgtatcct ttctctctcg 28140  
 gccccaagt tcattgatca tcttatcaaa cttttttaac ttatcaattt tgetttcatg 28200

caaaatctag taagagtaac ttcaaatata atccacgttt gatagtgaga ctcaagttta	28260
aatcatatc tatcttgcta tatgtaataa tcatgtttga ctgctgagtt tgatgggtca	28320
agactttcct aataaaataa aatgtgggta cggatgctta gttttgatgg gtgcaaatat	28380
aattgaaaaa ggcattgcac attaacaatg ttttacacgt ctaattttct cccctctgat	28440
tctcaagagg cagggacaaa cagatttcac atgccctttt ctgggataca aacatgggtt	28500
ctcactttct catggttttag ccttaaaaca tgattcatcg caatctgcc ttaccatttg	28560
ggaatgtgac tgaatgtttt ggtaactcag aatttgctac aatctgggtg taagtaatga	28620
ctattagatc ataattttgt tgagaataca atatttctat atattctaata actacagtgt	28680
ttctattttt ggtttcaaat ctcaaaaaca agttatacac aaattctttt gggagaaaat	28740
aaataataat aaaaaggcaa gctagcaatc aaactccga actaaagata cataacgagg	28800
tggtcacaga atagcttata cagtacaatt taagaatttg gtatacaaa tatgattttc	28860
aacaacgaac ctctctttacc caatatgttg tcacatttat ttgtaacctt ttaaaaactt	28920
ttgcgaatag ctcccttaat aaaaaatgcc gcatgattaa tcatcaacag gaaaaggcta	28980
gctcacttga tatcatgaaa agaaggcaag acagcaataa gacggtctcc ataaccaaca	29040
ggctctgtaa ataacaaca aaaaatagta aggaattcat cacaactatg gatgactgat	29100
tgagtgtagc tgccataatt gatggcctaa aatatgttta aacattgata atttgggtga	29160
gcattgacgt tgaacttcaa atattgcaa aggacggaaa tcgcaatgaa taaatcactg	29220
aaaagcata gcagaaatta agaccttaag taaacaatat tttccattc agtcaatagt	29280
cataactga ctgaaaaacc catgaaaacc gatatactgt aaattacaat gagctaaatt	29340
aatttatcat gaatatcctg ttactttcca tcatattgca aaaaggata caagattcaa	29400
gatccccagt gtttggtag attacaaaa agtcacatta tttccactt tgttttctgt	29460
tttaagatat tttttagttt ataacttata agagaaaaa agaatgttt tctcaaacct	29520
tattaagcct agatacttat gcgacccaaa tacgggggat acgggaaatt cttaaaaatt	29580
aagatacaac gcaactcaga tgcattaaca caaatataca cacacacaca aataaataga	29640
gagagacata catacactt ttaataaat gcacagtatt tattaagaga cattgattat	29700
cttaccactaa tacataacta tatcagtggc ccatgatcat tattcacaaa agcaatacct	29760
atgataataa caaaataaaa aaacaaaaa acagtgcact atgtttccat tactcatacc	29820
agaatatatt attttccaac atgctacaaa atcttcccc taacttgtgt caaaatgtca	29880
cagacatcat catttgtct actcaaaaga agaattgact ctatatgatt gcctagtagt	29940

aatatttagt atttactata gctttaaaga taaagctgta ttgaatttta ttctatttca	30000
aagtactgga gccattctca accataccaa tctacaaagt ttggaaaaa aaaatagagg	30060
atactctttg gaattggata agtacaagag tatcatatgc gtatcggtgt tgcatacaag	30120
tacagcatag atactttgtc atttttggag tatcaaggct tcacagctca aactaaacaa	30180
atcaaaccca acatcccac gttttaatat gataacagca tgctagccat aaagcaatta	30240
ggcatccatg atactagagt atcataaata caggtcagaa aatgctcaag tactaagtgt	30300
tccaagatgg taaaattcat ttttaagtct tatcactgtg actacgagaa tttagaatta	30360
ttaggattgc caaatcat ttgccatatt ttattgctag aggcacacta ttgctttaac	30420
tatttcaatt ttggatgaac agcacggcta tcatcactct ttctttcccc agaagccctg	30480
tattacttta gtaccatgta aataaatcta tacattttgg taacaggctca tagaaattat	30540
tatactcca tcttcaacaa gtacgttcaa cactttccca gccacatcag actgcaagat	30600
ttcaaaaatt aaactagtca gaagtagtaa atatttagga aggaaccaga attacagaaa	30660
cagaggcatc accctgatag gaagtccagt gccaaactga tccaaatacc ctatgacttg	30720
cccttctttg attacatcac cctagaattc aattaggaaa taaatatatt gaaaagaatt	30780
tgtagtcagt tcaatgaaag tgaggtcctc aaacaacttg atgcagcaac tgtatgatac	30840
aaaatatatt aataactaca ccagcagaaa aatataggtc aatctatatt tgggaaccaa	30900
ataatattta atttgtatct gatagactca agaaattata actaatttgg aagaaatgga	30960
tacctagtat tattaaaaca ccaaaacact gggcagatta tagtagctaa agaggaagaa	31020
gctaactagt caaagtgtca cactattcaa cactacaaag gaccaatccc cttttagaga	31080
gcctgacctt tctcacccaa gagctaccca agagaataca caccctctcc tccatatccc	31140
ctcccatata acacaatcct caccaactaa gcacctacct gacaattccc tccataccaa	31200
ctctctgtgc atcagggttg attctcttct ctttccaaga ctttgggctt ttgttttgac	31260
taagccaaat ttctatctgc tggcctggtc caacagtatc ttttacagac aagtttacia	31320
aatattcgta ttgttagaa tttattgata ttctattat gtccccactg tgtgcaaaac	31380
tttagaaact aatattacaa ttaacagttt ttgtgaatgc agcaaaacta aatatatatg	31440
atatagaaat caacaaaact gaaaaattat atgcaaagtt caattgaaaa gaaattgat	31500
tacccttttt gtggaataaa atataatgat aaactaggta ggttacagtt tggattttgtg	31560
atcaattgaa gatctagatg ctaattggctc ataactacaa tattttttgc agtgtctctg	31620
tgcctcacat cagtcacgtg tgtctatata actgttctct aaagtaata ttaaaataat	31680

ataaaaaata ttagaaattt aaattatatt tacatttttt aaatgtatta gtaagttttg	31740
tttatatcaa taatcaattt tttaagataa aaatttacia aaaagtaata cagtaataga	31800
aaaaataatta tcaatattgt ataagctgag actatcattg ctaattatta tcagcttttc	31860
ttttttatag catccagttt ttccagttaa atacttaaat ttattttcaa agccaacata	31920
tgcatgtcag caaggtcaca tcaggagact aggccgacaa gcaagcaggc tgatacatag	31980
actgcaacta ttagtttcag cactgcaaaa tgttagtgaa caacaaacac atgcacaaaa	32040
tccaattggtg gagttcataa tgagaatcca gatggaggga gtatcaagtg aattatgtga	32100
tctagtctat aacaaattga tgacataaaa acaccagaga agaggatgga tcttgacaag	32160
cttatcacat gtacctcttt acagatagga ggttgcttct tccctttcac tgtttcacct	32220
cttcggaata agccaaccta atgagaaaaga aagatctgtg atagctaact ctacatagct	32280
caagtgcagag ataattagtg aaagcagaag ttaaaacata tattattgat ggatgccata	32340
cogtgggaga tgttactaag acataggtat tgggtccaga agcctccaat gctgccaat	32400
tgggtgattt ctctttggaa acatttgcaa atgggttggt cttttcagga gatgattttg	32460
gtggtgatgg tggcaggcta ttgggtgctg attcatccat aggtttactt gggatagggtg	32520
gtggtgtagt cgttgaaatg tttagacaaa gaacctttgt tgcctcaatg tttcgcttaa	32580
tatgcatttc aaaaatccca acctataaca atataagcca aataaggttg tgtacaaggt	32640
catccagcta aaagcatgct tcttttgatg aaccagacag aaattacaga tttagagta	32700
tattaaagtt attaggcaca atgttgatg aaaacaaagc tgagctgaca caatttttta	32760
tgatcaaat tgcattacaa gaaaaactat atattcagat aggacattag gcacaatagc	32820
acctatttct caagtttcaa ttacaatgtg ctacctttt gataaaaaca taaataccat	32880
tacctattt aataatctat aagcctgtgt tatcttaagt gaattaaaa catcetttga	32940
cataagttca atacacagta catggcaggc tatagacata tcagacact aattacgaat	33000
tgcaagaac cctaccatat aacattttgc tttagaacat acacttaag ttttttcat	33060
cctgtttatt attatattta aaatactgca cttcaattt tacagaagca gtaacaaga	33120
gagaagaaga tcacctttac ttccagttca gcaatttcag tctcatcaca gacctctaat	33180
accaaagcct gataataagc aaaccagaca taagttgtgg ttatcagat tcacattgaa	33240
acatggctgt taacgaatgc aattatgcag tcttggtata aaaccagcta tcataagtaa	33300
aaaatacctt gtattttatt cttctataaa ctaacataag gtattttgat tcatgagcat	33360
acaattgcca attcccacca aactatcatt tttcactaat aagataagat tttcaggaaa	33420

ataaaaggaa agaatagcta tagacctcaa atccattagg aaaagtagca gtttgcagg	33480
gctttttctc caatgagcct tgggggggtgt tatctgaaga agctgcaaca tcataataaa	33540
ttgttaaggctc atggttttacg aatgttaaag taaaagttgt taaagggagc ctattctgat	33600
agccaactat aacaagctg tatatatcat ataacctta aacagacatt cacaattacc	33660
ataataataa ttagcaacag agtagaatga ttctctacta ttgtgtatct atgtaataaa	33720
cataatacct taaaaataa cagataaaca ataaaaaca catcttttgt ttgtaaggat	33780
tttctctctc ctctctctct tccaacaagt atataacggt tcggcaacat tatctttgat	33840
catgtatata aaaatatga attgcagccc aaaaattttg caactgggtta atgtatctca	33900
aaatcttaag aataaaaaa acttaagtt tattactaag aagataatta cactaagaag	33960
tatgaacaaa taagagtatc agacttttct cattcaaca acagccaat agaacaaaa	34020
cataaggata taattcattt cacaatcaat ataaccacat catcagaatt gtagtggttg	34080
atagcttcag ctgtttttgc agatgaaact agtgtgtctt tccccctcat gtgggaatta	34140
atgtgtctct gaccatattc caaatgtcg atgaaaagcc tacgtttaga gttccatctg	34200
gcattatgga tgggaagtac agcttgcttc tcaaggcagg ctgcacatg ggacatagtg	34260
cccatgggat ctgaaatcca agagtcatga ttaagcatct ggttgcaatg gtaaaatttg	34320
cccccttaaa aatataaaaa acattgaagt attacaaaa tattgtatta tgatttatga	34380
aaattgacc ttaagaagc cattaagaat ctttccaaa agaagtaagc ctatttcaaa	34440
aaggctcggg aaaagaacgg aacacattgt ttattttaac agaaatggat caagacaaga	34500
aaaaaaaaa aaccaacata acacaaatat ttctgacaag tgtttccaat caaaaatagt	34560
atacagccta tgggtaccac caaagttaat aaataataat ttaaaaaaa gaaaaacaca	34620
gtcccttggt tctactata tgaccaaca gaatgccaat tgcagtctct taggtgatgtg	34680
cagagcatcc ttggttata tatatatata tatatatata tatatatata tatatatata	34740
tatatatata tatatatata tatatatata ttgttagaaa catgaaggat acattcaact	34800
gccttcataa aaatgttatg ctggtgattc tgataaatat gtgacactta gcattagcat	34860
aaaataacct gtcaatgcta cctcattttt cgttaattt ctttgggtgt attcttcaac	34920
tgtttttcaa ttgattttt cctcatatat gtcacataaa ttaagcaaa taaaaccgaa	34980
aagcaagaga gcaagatcag aagtcgcaa cacacgagca gagatggcag tcggcaagac	35040
acgttcataa caaaaaaaaa atgcaggtag agatgaggag agagagagag ttacagtga	35100
aggaacgaat ggcaggcgag gattccatgg gaagaaatgg aaatggaaga atgggaggga	35160



aaaacaatgg agggaggagag gaacttatag agaagagaga atagccaagt tgagttaagc	35220
gaatgaggaa agaggtcgat aaattagttg cactgtgtct gctttgagat ttccgccctt	35280
ctaataccct tctccgttcc aatctaggaa cattgcctcg ctaacgtgcg ccggtgtgtg	35340
actagtgtct cctctctccc tcttacagtc ttacgtggga cccacccctc caggcaggta	35400
ggtttcatgg acagccatac aatgaatagt tcaaaaagtc taatttagta gtttcttgtt	35460
actataattt ttttatgcag tccacactaa taaaaaatta gatggttgga aaacaatct	35520
tattacaagt tttataggta aacttgaaaa actctatgtt ataagacctt tttctcactt	35580
tggtagtagt ctcttatcca agttagataa ttcttcttat cttaataata atattttttt	35640
tatagtgata catggatgtt atttagtggg attttattat cctctctccc acctactctt	35700
tcattatagt aatgcattct tcaagagtc aaaaatatatt tcattacttc caagaataaa	35760
ccttttaatt ttggatagat ttatttttta gtcttttaatt ttatttattt tttagattta	35820
atttggctct tcagtttttc agaattcaat ttaattctct aattttttta atcgatcaaa	35880
tttggttttt caatctaaat tataagaac tatattttgt gatggtttaa aatcgccatt	35940
aagtgttctt aagctaccac aaaagcaca ttccaaaaa aataaattga ttttaaaat	36000
tataagatca aattgaatca attttaaaaa ttaaaatatt aaattgaaaa aaaaataaat	36060
gatcaaatg aacataaata ataaatttga ggattaaaaa actaatttaa cctttaattt	36120
tttctcactt atattaatat taaaaaatta tattgatatt cctaataact ccttatctca	36180
attaaaaatt ccaaaaatta attctagcat ctccaaacac tactcaccat gaaagtccat	36240
cacaaccatc tttctttctc ttttctctac atcatgtttt cgcttcgcaa actttattgt	36300
gttcctagtc tttagcgtct gataatcttc cacaagtatt gaactataac acttattgtg	36360
cttgccaccgt taatagctaa caccaaatga gacgtgtcac ttgactttta tatcactaag	36420
aaaatttcaa cacattgac cagtattagc tccatcttgc tttaacactt gtttgactag	36480
tcacttaagt gcaacaacca actttgatat cattgttgga aaaataaac ttattagaag	36540
tttcctagac aaacacgaga aactctttcc attacaagac ttctctatt acttgggtat	36600
ggtggtgact tcctttataa tgggtgtgaa tagctccatt tataaatgtt atttagtgag	36660
ttttaattat atcatctctc tatccacatt ttcattacgc tagtaggatt ctccaaaat	36720
caagttacat tccattttac gtcacctctt aattttttgc tcagttgctt taatatttgg	36780
aaatttggat tggttttcac aacatatact agatataact tttaatgtaa ttcaaaaata	36840
caattcttga taaattgatt ttcacaacat attcatatat actcaaaatt taaagataag	36900

tactacattt tatcaggtag gtcaacacat tttaacctcc cgatcataga gtgatcaaga	36960
ggaaaaaaag aatgaatgaa aggaggagta aggtgaagga aagtaatgaa aagaatgaga	37020
caacttttaa aaaattaaaa ttaaggataa taaatttatt tattgaaata aggggtttatt	37080
ttaatcaaat aaccaacttt tttgtttttt tagtatgttt gtctaaatta ttatttttaa	37140
aaaataactt tctgtttatt ttaagaaaca aaacaaatct tatttgcttt tttaaaaaat	37200
acttatataa aaatatattat tttttttttt ttttaagttt aaacaaactc atcctaatat	37260
gaaccagaaa accttagttt ttgttaacaa aaatgagtta aaataatttt ggattcattt	37320
taaaactatt tctttgtctt tttagtcaat caaatgattt attgataaaa aaatatcaa	37380
atatttttgt tttcaaattt aagtggagg tgataaaaac acttttttta ttggatagaa	37440
tttattaaaa tttacaaaaa tcatgagtga agttagaatg atacatacat attttgtcat	37500
ttccaataat ttttagtgaa aataaattgt attaaaaagt gtgttgctat ttttagcact	37560
gagagtcagt aacatggatt tgcgccagga gtgataattt gtggaatcaa gtgagggaga	37620
aactcatttt tcaatttaac tttaaaaacc aaaactaaaa aacttacaac tatacattgt	37680
attaattagc atgtgtttta tatatatata tatatatata tatatatata tatatatata	37740
tatatattgag tatggaagga gtactctatt caatgagatg aatatgtgtt aacaaaaaga	37800
ttgattagtc gattaagaaa gaagagagat tcaattcttc ttactactaa aatctaataa	37860
actgataatt aacatttggt aataataaaa aaatgaaatg gacatgcaat taattaggcc	37920
aatgataaag aatatattta aaacaaattg ttaaatcacag tgtgtttgac aatgatatat	37980
aatcggtgcc atggatcata tcaacctgaa actaatgaaa ggatcacaga tcaattctat	38040
cttccaatta aggaatcaca ggtttaaaca atagtagtag ttaatttgca catcactact	38100
ctggaggcaa ggctaagcaa cgtcgagatg gactatttct cgaaactcca acctctagt	38160
ttcactgtag aatgtcacac attttgttag accaatatgt tagccatata caacccctta	38220
tttccatttc cgtgtttttc cctatggctc cttgtttcac ttccaccac actttttttc	38280
cataattcat acctgcgaat ataaccatc cttccttgtt gaaaggtcag taagcactgg	38340
agtatcgatt ttgaaacct gactttgggt taagcattgc aaacttcgaa ggctttgatg	38400
ataaaacttc ttggcatca acgctgtcat gaagaaacca attcaaattt gcactgtatc	38460
cataaactaa cggttgagga ctctctccag ctgacacctt aagtgcgaag ggattcgatg	38520
ctcctaatga agctctacag cttttgtcaa caccaatttc ttttcgtggg ttctctgagg	38580
atttgccaac tgataggctc aacacttggc cttggattgt gggtgtaagt tgcaatgata	38640

ttctggaagg	aaaatgtttt	tcagctgcac	cagccctctg	ttgtttcata	aggaactcat	38700	
tcaccaattc	gataacatca	cacctgtcca	catggaatat	tgtaaattat	attggagatt	38760	
aagaatttac	gttaaaaaat	aagagtgaat	ttgtttaaac	ttataaaaa	taatttaaaa	38820	
atagatat	ttttatataa	agggtttttt	aagaaataaa	taagattttt	ttatatatta	38880	
aaataaaaa	attattttgt	taagaaaaaa	taatttataa	aatgttttta	aaaaataaaa	38940	
aattacttat	tttattaaaa	tcaatattta	ttttaacaaa	taactttaaa	taaattgacc	39000	
ctaaattgta	cacaattaat	taagaaaaac	agttactcaa	catatatata	tatatatata	39060	
tatatattac	aatcttattt	ccttaggtga	gatgaattag	atagatcaca	tatcatcatg	39120	
actcgattaa	taatctcctt	aatacagtga	agtaaataata	taaagataat	gcgattcata	39180	
aaaagagtga	ttttgatcac	ctttgttata	aaaagaaaaa	gaacgatgct	tttttttttt	39240	
ttgtaaatat	aaaattttaag	tgacttttag	atatttccaa	tctaaatcca	aaatatgtag	39300	
ctataagaaa	ttgagtttaa	gttgaaacaa	agtatatcat	agaaggaaga	atttagagaa	39360	
aagaaaagaa	aaaagtacat	accacacctt	gtaacaattt	caagctaattg	agcccaaat	39420	
tgtacaaatg	atgaattttt	ttttcctaaa	cttctaataa	caacatggcg	ataagctgaa	39480	
taatggataa	tttattttgc	ttgttaagat	gcttatttat	tggtcttata	tgtattatct	39540	
cttaattggga	tattaatggg	ataatattag	tggttaatta	attagtgggt	accttacctt	39600	
gtcaatgtcc	accatgat	ttt	cagccactc	gtctttgatt	aagactttgt	gattgaattg	39660
aagaacgcct	tccacatggt	gatatctgag	ggagaattga	agcctctcgc	tagctggact	39720	
gctttttgaa	ttctcaacat	ccatgacat	gaggtgacaa	tacagctgca	ccttccacaa	39780	
cccaattggt	gaaaatgcat	aagaaaacaa	aggacacggt	gttctgaagg	gggtgttatg	39840	
tgcttccaag	ttccccaacc	cagttactga	ccgcaagggt	caatgaacgc	atccacagct	39900	
cttccaagtt	agaccttaag	agtttcataa	tcatatgtga	tgcttgctta	gactgaaaac	39960	
ctatcaagtg	atcttttcaag	ttactaatgc	accagaaagc	aaaatctgca	gcaggagctt	40020	
catatatgta	aactaggaac	aagagggtga	agaagaaagc	tctgaaaggt	tgaaaacgct	40080	
tgaaagggtg	tcggaagcag	tggaagtcag	gtttggaaat	ctgatgaagg	tagtgctggt	40140	
tttgtgtggg	ccataaagaa	ggatggcctg	aatgaaatta	gcaaagacat	tgtaaatggt	40200	
ttcttcatca	tctattaact	tattattggt	gttttttagta	gtgctgctaa	ttggcttgaa	40260	
agggttttag	gtctaaagat	ggatgggaat	gttgcagtct	gcaataatgc	caagatagag	40320	
gtttgaggaa	tgactattct	ttttgtgtac	ggtgagatta	agacgtgggt	ggcatgagct	40380	

tgaagaagaa catatgttta aggacatgga acttgtttcc cattctgaaa ttggtggaag	40440
gttctgaatc cagcaaaaca catcaagaaa gttgttagcc atggatcgaa tgaagcaact	40500
taattaataa aactctctct ctctctatct ctctaattcg gttgcattca ggtgtggcct	40560
cacatttatt tgtagactct tacataatgc tatgttatgt actgcaatta gcaataactc	40620
tttctagtgg agaaataata attaaaaaag tggactgatt ggtacgacca ttagtttaat	40680
tagctccatg gagaaaagca agataaaatt gctaattatt ggttaagaaa ataattgcac	40740
cagatatatt atataaaatg tcaaaaacgc attccgtaca ttataataa tattatatac	40800
gtcatattta catcattttt tatcctgtgt tatctcaaaa aagtgttaaa atagagagag	40860
tatatatcat atcatataat atgtaagtgt ttattagtgt aaaaaatag cttgagagta	40920
atgtgatttg tcatgtgcta ataaaaatc attttgaatg ctcttttacc cacatatatt	40980
aattgttaat gattgaagtt tattattatt attataatat ccttttaacg atgaaagtgt	41040
gttttaaaaa aatatagatt taagatgtgt ttggaggaa ttatttatat cttatctgaa	41100
cttattttat ggcatacgtg taagtattta agaaaactta taaaattata gtttatgatt	41160
tatttataaa ttgttttcaa cttattttta taaaattttc aaaataactt ataagaacaa	41220
attaaatttt ttatatgaaa ataatttaac cttattttct ttcaattat aaaaaacaat	41280
ttacaataaa aagcttatat atatgatata cacttttaag ttgttaagta agctatctaa	41340
aaaaggccgt acagtggttc tttaatgaac tatcgatcgg gaatgttata tatggaaata	41400
tataacttg agtgaatata ggctcgatta ctccatagta cagtcaata attattagta	41460
aacgaattat acgtttaatt tgtatctata tatcttttgt tgataattga tgtaatttca	41520
attttaattt accaaagaga gtttagacca cagcgagcat cagtgcctc attagtcat	41580
agtacttacc accgacatct ttttgtttgt aaaaggacca ctgattcatt tacctacata	41640
tataatatac aatatgtatg tatacaaaaa tcatagtaag gtttaaatgt aatgcttcat	41700
gaataagata ttctgtgtta cagattaaga ttctgtgatg ataaaatggt tgttattatt	41760
agagttaaco ggcaatttgt tcatattgag tctcattaat taccttcttt tctcatgttt	41820
tgttgacac gagagtgacg atcctaccga gatagataag gatatatagt gatacaaat	41880
tgagataaaa agctctttgc acagtcaatt atgattaaga aaaatatcaa atcagtttta	41940
cagaccctag ctcattaggc agagataaatt acatgcacgt aaagaaaaaa ttattgagtc	42000
actaaaattg ggatagcgag gaatttgagt aatttgaact aagtcataag tttaaatcgt	42060
atcgttaaaa aaatgtagt ttttgttact cttttaaatg ctagtatttt tattttgaaa	42120

ggttttaaat accagtatta ttccactaat aacctgcctt tatttcttta tataaagcct 42180  
 tctcttgatg aaaatagaat actaattaaa taatcgagaa aaaaaagata caaatggaga 42240  
 acaaaattatc atgaaaaaat tacacattag aaaatataca tgttttagca ttgaaaaata 42300  
 caatgggtcaa ttataaacca aagaggccta actaagtcta atgtttacgc caccaaattt 42360  
 ttggttgata acgttttaaa gtaatatgta gatggtctct tcaaaagaaa tttctgtcca 42420  
 tattattcag gtttcaaatt ttgtttgtaa gacgaggaat ttgggatctt gatgataaga 42480  
 ataagacagg gtgaataagt tcatttaatt aagatggaac gtgcgagttt aacttgagtt 42540  
 acgtgtgaagg ttccataatc aaggtgtacat atgtatatgt attagggatg attaatgata 42600  
 ttagctatca aatttaataa aatgtatat ttatattatt tttttatcaa cagtaaattt 42660  
 tgtaatttta acagttgaa tttaaagttt cataaaataa attaaacccc acattatttc 42720  
 aaaaagtaat taatactttg ttactacact cttaattata tgcataatgc atcatatttt 42780  
 gtaataaaaa ctttatattt acacacgtat gaccattgat taacctacac tgtggcaagt 42840  
 acacctcat tttctaact ttataaataa aagttttttt aaacagaaaa ttatgataaa 42900  
 atcttataat ttttatattt atttcattta tttttatata tttattataa attcctatat 42960  
 ttatatattt gaaccactc attttacttt ttataattta tccacattga ttcaagttcc 43020  
 aaatctacac ccactgagtg cataaattaa ttggaatatg ttttaactta atcaagggtc 43080  
 ttgagttcaa attttgaata tggagttact tacatattta aaggaagagt ttgttatcca 43140  
 taatggttcc ataagactct cgaacaaaaa tacttccaat aaaatataca tgtggtttat 43200  
 aaaaaacaat tccatcaaaa ttttccaaaa ataatacaa aaggatacaa atattttttt 43260  
 taaaaaata attcatttat tttgaatata tgacttttat atatatatat atatatatat 43320  
 atatatatat atcaaccggg acatagtaat tcaagactac ttaatgttgt tcaccctga 43380  
 tacatgttaa ttcaatatta cacaatcatt atcaaattha attttagaaa atttaattat 43440  
 tttctccatt agcatatagt catttttatt ggaaaaataa ttgatgaaa catatcatac 43500  
 taattaaagg ataaacatta taatttataa aagatttcaa ctatatccat taattgtaaa 43560  
 gaaaattttc aattgagaat cgaagttaat aattatcaaa ataattcttg cttttattta 43620  
 tgaaaatata ttgtgtgatt cttaattatt ttcgtaata tataaaaatg aatatcatca 43680  
 tatattttga agtaacttaa aatatattta atcctaaggt tctacatgct tgaacaaacg 43740  
 tcttcacac aaatctttgt agaaaaagta aataagacac taccaaaaaa aaaaaaat 43800  
 caccaccact acaataaaaa aaggtacgca aaaagagagg ttacactatt accaccctac 43860

100

gtcatcataa aaccattacc atttaccact cttttgaacg ttaatgtagc ctaaatctta 45374

tatccagaga accagaccct gtttaaattt cttttttaaa atgtttctga taaatttctc 45434

ttgcaagtgt ctacag a acc cag tta gct cct tcc tca cca cgt gac act 45483  
 Thr Gln Leu Ala Pro Ser Ser Pro Arg Asp Thr  
 55 60

tca gtg aaa ctt gga gat gcc agc ag gtttatttca gccagggtct 45529  
 Ser Val Lys Leu Gly Asp Ala Ser Ser  
 65 70

ttgtctctca gggcaattca ttaattttaaa aaataacatt tttttatata tattcatcag 45589

tgcacgagga ggagggatag tatgtatcac acttttttaac tcaactttcta ttgttttctg 45649

ttagttgaaa ttcaaatatc cctcactaat ttgagactga aacatttcac caaaaaaaaa 45709

aaaaattgag gatggaaactt tcttttttag ttgatcataa attttttctt ctaaaatata 45769

taatgtggat acatattttt tgagattgaa acctaacaaa tgataaataa gactcactta 45829

tttagtgaga catcacatgaa ttccagagaa tatttttcta tataggttat tagcatttct 45889

tttaatattt tttttattat tgtcttattt ttaaaaagtt ggcattcttt ttaaaattga 45949

cttttttgag atattgaact attttaataa taataataaa attaagttat atagtgtatt 46009

aaaagaata agataaaatg tgtttttaaat ttctcaagac tttagtcaaa attagtttca 46069

gtctctctca ttaaaaatgt gttttaattc tcataatttt aaaaatatgg tgaatttcat 46129

ttttaatctt gaacagtctt ttaattttga cttaattaaa ttcaacatat ttcagaaaca 46189

cgggaaccaa aaccaccatt tttagaatcc aagactataa gatcttaatg acgtaaaaa 46249

caattttacc gtgagaatat taaagctagt agtattgctt ttcagtgtgt ttctacggc 46309

acattgttgt gtgtggaagt ggaagctaga aaacaaaggc agcagaagaa gtatggctct 46369

acaaagtgtg tagtagtgaa gaagaaatag ccgttggtgg tggagaggcg cggggttgca 46429

ataaagaac agcgcgccat gatcctataa taaacctgtt caacaaaaac aagtatgctt 46489

catgaatagt tactatttac aaggaaaact agccgttact cactttttct tctttttttt 46549

ttttgtaaca aattctgaac cctgcatggt cattctctct ctctcacgct cgcaaccgcg 46609

gcgcgcacct acacttcttt tatgtcatca cgtgtctctt ctactctctc ctctctctca 46669

ctacaaaaac cattcttcaa cttgcaacac acgcacacac acacactcac acacactggt 46729

ttttgtttcc actaaatcaa aacctcttat ctcttactct cattacattc attcttttga 46789

ttttcgttat ggttagtagca gtggagaaaa ccaacctcac ttacacaatca caatgcttca 46849

accgtgtttc tgacaagaag aaagaaagat gcaagacaca catgaacaac gtttaacctat 46909

gtgtgtttttt gttttcttta tgtgtgtgga g-c ctt gtt gtg ctc ccc tca tgc Leu Val Val Leu Pro Ser Cys 75	46962
gtg agg cca gtt ttg tgt gaa gat gaa ggt tgg gat gga gtg gtt gtg Val Arg Pro Val Leu Cys Glu Asp Glu Gly Trp Asp Gly Val Val Val 80 85 90	47010
aca gca tca aac ctc tta gca ctt gaa gct ttc aag caa gag ttg gtt Thr Ala Ser Asn Leu Leu Ala Leu Glu Ala Phe Lys Gln Glu Leu Val 95 100 105 110	47058
gat cca gaa ggg ttc ttg cgg agc tgg aat gac agt ggc tat gga gct Asp Pro Glu Gly Phe Leu Arg Ser Trp Asn Asp Ser Gly Tyr Gly Ala 115 120 125	47106
tgt tcc gga ggt tgg gtt gga atc aag tgt gct aag gga cag gtt att Cys Ser Gly Gly Trp Val Gly Ile Lys Cys Ala Lys Gly Gln Val Ile 130 135 140	47154
gtg atc cag ctt cct tgg aag ggt ttg agg ggt cga atc acc gac aaa Val Ile Gln Leu Pro Trp Lys Gly Leu Arg Gly Arg Ile Thr Asp Lys 145 150 155	47202
att ggt caa ctt caa ggc ctc agg aag ctt agt ctt cat gat aac caa Ile Gly Gln Leu Gln Gly Leu Arg Lys Leu Ser Leu His Asp Asn Gln 160 165 170	47250
att ggt ggt tca atc cct tca act ttg gga ctt ctt ccc aac ctt aga Ile Gly Gly Ser Ile Pro Ser Thr Leu Gly Leu Leu Pro Asn Leu Arg 175 180 185 190	47298
ggg gtt cag tta ttc aac aat agg ctt aca ggt tcc ata cct ctt tct Gly Val Gln Leu Phe Asn Asn Arg Leu Thr Gly Ser Ile Pro Leu Ser 195 200 205	47346
tta ggt ttc tgc cct ttg ctt cag tct ctt gac ctc agc aac aac ttg Leu Gly Phe Cys Pro Leu Leu Gln Ser Leu Asp Leu Ser Asn Asn Leu 210 215 220	47394
ctc aca gga gca atc cct tat agt ctt gct aat tcc act aag ctt tat Leu Thr Gly Ala Ile Pro Tyr Ser Leu Ala Asn Ser Thr Lys Leu Tyr 225 230 235	47442
tgg ctt aac ttg agt ttc aac tcc ttc tct ggt cct tta cca gct agc Trp Leu Asn Leu Ser Phe Asn Ser Phe Ser Gly Pro Leu Pro Ala Ser 240 245 250	47490
cta act cac tca ttt tct ctc act ttt ctt tct ctt caa aat aac aat Leu Thr His Ser Phe Ser Leu Thr Phe Leu Ser Leu Gln Asn Asn Asn 255 260 265 270	47538
ctt tct ggc tcc ctt cct aac tct tgg ggt ggg aat tcc aag aat ggc Leu Ser Gly Ser Leu Pro Asn Ser Trp Gly Gly Asn Ser Lys Asn Gly 275 280 285	47586
ttc ttt agg ctt caa aat ttg atc cta gat cat aac ttt ttc act ggt Phe Phe Arg Leu Gln Asn Leu Ile Leu Asp His Asn Phe Phe Thr Gly	47634



290	295	300	
gac gtt cct gct tct ttg ggt agc tta aga gag ctc aat gag att tcc Asp Val Pro Ala Ser Leu Gly Ser Leu Arg Glu Leu Asn Glu Ile Ser 305 310 315			47682
ctt agt cat aat aag ttt agt gga gct ata cca aat gaa ata gga acc Leu Ser His Asn Lys Phe Ser Gly Ala Ile Pro Asn Glu Ile Gly Thr 320 325 330			47730
ctt tct agg ctt aag aca ctt gac att tct aat aat gcc ttg aat ggg Leu Ser Arg Leu Lys Thr Leu Asp Ile Ser Asn Asn Ala Leu Asn Gly 335 340 345 350			47778
aac ttg cct gct acc cta tct aat tta tcc tca ctt aca ctg ctg aat Asn Leu Pro Ala Thr Leu Ser Asn Leu Ser Ser Leu Thr Leu Leu Asn 355 360 365			47826
gca gag aac aac ctc ctt gac aat caa att cct caa agt tta ggt aga Ala Glu Asn Asn Leu Leu Asp Asn Gln Ile Pro Gln Ser Leu Gly Arg 370 375 380			47874
ttg cgt aat ctt tct gtt ctg att ttg agt aga aac caa ttt agt gga Leu Arg Asn Leu Ser Val Leu Ile Leu Ser Arg Asn Gln Phe Ser Gly 385 390 395			47922
cat att cct tca agt att gca aac att tcc tcg ctt agg cag ctt gat His Ile Pro Ser Ser Ile Ala Asn Ile Ser Ser Leu Arg Gln Leu Asp 400 405 410			47970
ttg tca ctg aat aat ttc agt gga gaa att cca gtc tcc ttt gac agt Leu Ser Leu Asn Asn Phe Ser Gly Glu Ile Pro Val Ser Phe Asp Ser 415 420 425 430			48018
cag cgc agt cta aat ctc ttc aat gtt tcc tac aat agc ctc tca ggt Gln Arg Ser Leu Asn Leu Phe Asn Val Ser Tyr Asn Ser Leu Ser Gly 435 440 445			48066
tct gtc ccc cct ctg ctt gcc aag aaa ttt aac tca agc tca ttt gtg Ser Val Pro Pro Leu Leu Ala Lys Phe Asn Ser Ser Ser Phe Val 450 455 460			48114
gga aat att caa cta tgt gga tac agc cct tca acc coa tgt ctt tcc Gly Asn Ile Gln Leu Cys Gly Tyr Ser Pro Ser Thr Pro Cys Leu Ser 465 470 475			48162
caa gct cca tca caa gga gtc att gcc cca cct cct gaa gtg tca aaa Gln Ala Pro Ser Gln Gly Val Ile Ala Pro Pro Glu Val Ser Lys 480 485 490			48210
cat cac cat cat agg aag cta agc acc aaa gac ata att ctc ata gta His His His His Arg Lys Leu Ser Thr Lys Asp Ile Ile Leu Ile Val 495 500 505 510			48258
gca gga gtt ctc ctc gta gtc ctg att ata ctt tgt tgt gtc ctg ctt Ala Gly Val Leu Leu Val Val Leu Ile Ile Leu Cys Cys Val Leu Leu 515 520 525			48306

ttc tgc ctg atc aga aag aga tca aca tct aag gcc ggg aac ggc caa Phe Cys Leu Ile Arg Lys Arg Ser Thr Ser Lys Ala Gly Asn Gly Gln 530 535 540	48354
gcc acc gag ggt aga gcg gcc act atg agg aca gaa aaa gga gtc cct Ala Thr Glu Gly Arg Ala Ala Thr Met Arg Thr Glu Lys Gly Val Pro 545 550 555	48402
cca gtt gct ggt ggt gat gtt gaa gca ggt ggg gag gct gga ggg aaa Pro Val Ala Gly Gly Asp Val Glu Ala Gly Gly Glu Ala Gly Gly Lys 560 565 570	48450
cta gtc cat ttt gat gga cca atg gct ttt aca gct gat gat ctc ttg Leu Val His Phe Asp Gly Pro Met Ala Phe Thr Ala Asp Asp Leu Leu 575 580 585 590	48498
tgt gca aca gct gag atc atg gga aag agc acc tat gga act gtt tat Cys Ala Thr Ala Glu Ile Met Gly Lys Ser Thr Tyr Gly Thr Val Tyr 595 600 605	48546
aag gct att ttg gag gat gga agt caa gtt gca gta aag aga ttg agg Lys Ala Ile Leu Glu Asp Gly Ser Gln Val Ala Val Lys Arg Leu Arg 610 615 620	48594
gaa aag atc act aaa ggt cat aga gaa ttt gaa tca gaa gtc agt gtt Glu Lys Ile Thr Lys Gly His Arg Glu Phe Glu Ser Glu Val Ser Val 625 630 635	48642
cta gga aaa att aga cac ccc aat gtt ttg gct ctg aga gcc tat tac Leu Gly Lys Ile Arg His Pro Asn Val Leu Ala Leu Arg Ala Tyr Tyr 640 645 650	48690
ttg gga ccc aaa ggg gaa aag ctt ctg gtt ttt gat tac atg tct aaa Leu Gly Pro Lys Gly Glu Lys Leu Leu Val Phe Asp Tyr Met Ser Lys 655 660 665 670	48738
gga agt ctt gct tct ttc cta cat g gtaagtcttcg tgtgctgttc Gly Ser Leu Ala Ser Phe Leu His 675	48783
tttcattaag tgttgtgtgt gctgttcttt aattataatt tggagtttta ccttagtaat	48843
ctgtataatt ctaatcgagg aacagtacaa acaaaaacac ctaaggaaca ctatagcacc	48903
taaggaacaa caccttagct ttaatatacc atatcaataa gtgaattatt tctctgttca	48963
tcttgatgca g gt ggt gga act gaa aca ttc att gat tgg cca aca aga Gly Gly Gly Thr Glu Thr Phe Ile Asp Trp Pro Thr Arg 680 685 690	49012
atg aaa ata gca caa gac ttg gcc cgt ggc ttg ttc tgc ctt cat tcc Met Lys Ile Ala Gln Asp Leu Ala Arg Gly Leu Phe Cys Leu His Ser 695 700 705	49060
cag gag aac atc ata cat ggg aac ctc aca tcc agc aat gtg ttg ctt Gln Glu Asn Ile Ile His Gly Asn Leu Thr Ser Ser Asn Val Leu Leu 710 715 720	49108

gat gag aat aca aat gct aaa att gca gat ttt ggt ctt tct cgg ttg	49156
Asp Glu Asn Thr Asn Ala Lys Ile Ala Asp Phe Gly Leu Ser Arg Leu	
725 730 735	
atg tca act gct gct aat tcc aac gtg ata gct aca gct gga gca ttg	49204
Met Ser Thr Ala Ala Asn Ser Asn Val Ile Ala Thr Ala Gly Ala Leu	
740 745 750 755	
gga tac cgg gca cct gag ctc tca aag ctc aag aaa gca aac act aaa	49252
Gly Tyr Arg Ala Pro Glu Leu Ser Lys Leu Lys Lys Ala Asn Thr Lys	
760 765 770	
act gat atc tac agt ctt ggt gtt atc ttg tta gaa ctc cta acg agg	49300
Thr Asp Ile Tyr Ser Leu Gly Val Ile Leu Leu Glu Leu Thr Arg	
775 780 785	
aaa tca cct ggg gtg tct atg aat gga cta gat ttg cct cag tgg gtt	49348
Lys Ser Pro Gly Val Ser Met Asn Gly Leu Asp Leu Pro Gln Trp Val	
790 795 800	
gcc tca gtt gtc aaa gag gag tgg aca aat gag gtt ttt gat gca gac	49396
Ala Ser Val Val Lys Glu Glu Trp Thr Asn Glu Val Phe Asp Ala Asp	
805 810 815	
ttg atg aga gat gca tcc aca gtt ggc gac gag ttg cta aac acg ttg	49444
Leu Met Arg Asp Ala Ser Thr Val Gly Asp Glu Leu Leu Asn Thr Leu	
820 825 830 835	
aag ctc gct ttg cac tgt gtt gat cct tct cca tca gca cga cca gaa	49492
Lys Leu Ala Leu His Cys Val Asp Pro Ser Pro Ser Ala Arg Pro Glu	
840 845 850	
ggt cat caa gtt ctc cag cag ctg gaa gag att aga cca gag aga tca	49540
Val His Gln Val Leu Gln Gln Leu Glu Glu Ile Arg Pro Glu Arg Ser	
855 860 865	
gtc aca gcc agt ccc ggg gac gat atc gta tag cacaaatttt gcattgattt	49593
Val Thr Ala Ser Pro Gly Asp Ile Val	
870 875	
ttttgtgcca aatgtagtag gctactata tatatgttct atgattcttt cattcttata	49653
ttatttttgc ctgtttgaat gcttgaatt gtacatactc atactacaat aagggtgtagt	49713
ctctggttaat ttacacctca cctcaaagct ggggtgtaat tctgtttcct ccaaggcaca	49773
taatagtga aaatagttct caggagcatt cattgtttat tctgcaagat tctctttcac	49833
ggctgctato ttctatgcat gccctgccca taaatgcatt atgaagaatt gtaacggctg	49893
tgttttttgga cttcttcaaa aagtttatgt tattgcccagg tgtatatatc aacatgtttt	49953
aaagattttt aaacaatcag gtttttagatg tgggtttgca tgcattgagat tggactagtg	50013
cgcttgatgt agtataaaat ataaattgtc caatcagcac cctctacatg tccaaataat	50073
gggccttatg aaacttaatt ttttaattac aaactacagt aatctttttg aataaagatt	50133

tacaaattac aacagacatg tgaagtcgtc atctttcatt gccattctt tcaagtttac	50193
tactattatt ttctgtcaag cattccacat tcacatctga taactatgac agcatcttcc	50253
aagataatga cttccaagti ccaacactgg ctctgtacat ttgaaactaat tttatatcat	50313
ttatctattg tgattgaaat ataaaattga agtgaatgga acaatacaaa tcacatcttg	50373
aattaaaata tctaacaact ggaacaaata agaggcccag aaaaaaggga taaataacgg	50433
ataacaagac agaagaaaa aaacccaac ataattccaa cttcaaaatt cactcaataa	50493
aaagtttaac atgtaaattt acttggaac aaactcata accaataata ataataataa	50553
aagaaatcag ttttatgca ttaattggg atgctctgct tgtatgcmaa tggcacaacc	50613
ttaccctcaa gattgcaaaa cacagatgag taacagatgc aatgtgaatc aataaaaagt	50673
attgttcgct tgttgatgac acaaccttac tcataaaaaa tgcattgttg atggctagca	50733
ttgttgcaag gtattcatat agttttattt gcaacataga gaaaaataca ctcataca	50793
ccaggaaatg gttctgttca aaatcacgga ttataaaaag ttattatcta aatgttaca	50853
gcttaagaag atctatcatt gtgaaagtct cttatgcaca ttaabattac aagcttaaga	50913
atgctatata aatgtttgag gttttgatat ttaacttttt atgatatgct ttgattta	50973
agttgcaaat tgcccatttt ctcatgtcag ttactcgtat tctccataa ataaataagg	51033
cttctctctg ctcaatttat ttacttcta aagcaacaat ttctttcttt ctcatttttt	51093
ttgtaccctc tgatcagatc gtagtcogac cctcaagcct tagcctctac acactttgtg	51153
ctcgaggctc gatgatgtga tatgcttctg accggactgt agtctaattc ggtggtggat	51213
cccaatctga tctgataacc tcacagtat tgtgcacatt ataacactga cgtagggtta	51273
accatgtacc gagatccttg ggcacagcat cttgatagc tcaattggcc tcgattcctg	51333
aatatcggag atgggagagt aggcaggctg caataaaagg tcataccatt aacgtaggat	51393
ggtagggtta gtaataatgc ccttgacttg atgctctcaa ctaggtcacc ttgacatacg	51453
acaaagcctt ttctaaaca atttttata ttgaaatttg gaaatatgta tgcattgaaa	51513
ttcacacaaa tacttttcta actcatatcc aaatttaaga tcagcctaaa agoccaaatt	51573
aggggatgca aagaattcta tcgttaaatc aaacacatca gtcaaaaaag aaacataca	51633
taggcacaat caactgatca agtataacag tatttcagat attgttagcg actagaaggt	51693
ttgtatttaa tttgcaaaat ttaatatgaa tactactggt tctattttat ttataaactg	51753
gtattagtta tgaactttta aaatggggaa tatttttagca aaaaataaag gttaaaaat	51813
gtctatgatt cttaatgaat attccaattt tatgtttgct ttttagtata aaaaaattcc	51873

gtttttgttc cttgataaaa aaggaaattt gtttttgggt ttaaacactt tttttagtcc	51933
ctattaaatt acgaatttcc tatttattct ttgataattt ttgtttattt attagttcct	51993
tcaaaaatta ctaataataa atatttttta acaggaaata aatacaaaat tctctaattt	52053
attaggaatt aaaaaaatg ccaaggacaa aaaaattatt attttattaa gaattcaatg	52113
caacaaaaaa tttacctaaa agcaataaaa aaatttgagt atttcttagg aattaaaaat	52173
attttaataa aataaaaaa aagatccaaa tgatagtgtg ataaccgaag aggaatgtct	52233
ttcaaccact gcctgaccgc caccactgcc aacagcctag tatcaaccga atccacatat	52293
accaacaatc ttcacagaaa cacttctaag ttggtgtcga agagacaata tctcatgggt	52353
agatcaaatt aagagtgtct ccaataacaa aatcgggagc atttgactaa caaacagtta	52413
tgtgcattgg atgttctacc atagtacatt gctttatgtg aaagtctttt aattattcaa	52473
tattgacatg ttcttatata tatatatata tatgagggat tgtattatct ctgaaaaaag	52533
attttatcat aaaatcataa tgatttctca taatgtatct ttacatttta aagttagata	52593
aataaaattg attttaaatt gttagatata attaaaatac ataattaata tgacttttaa	52653
caaattgata tataaacact taaaaaaaag ttctcatgag tacggtgtgt attgttggta	52713
caaaaaaat ttataactat aactaattaa aattattata aataataaaa ttaataaaaa	52773
ttactataat aatctgtaat tagattattg taaaattgtt ttataataata aatatacagt	52833
cttttttctt taagaaaaat tgctagacca agcaatatgg accatgtgct ttctgaaat	52893
atataacaca aaatttccat taagtttttt tgcacctata agctacatcc gctacgtact	52953
gcatgtggag cctcatgagt gtgaggatct tccacaggtc actagtttga catctgaaag	53013
ctcctctgtg aaaactgtga aacaaataac aagcttgagc tgggtgtcga tttagtgtta	53073
ctagctatcc catgtaataa atatataaat cttgaatcac aaggaaatgat gcaatatatg	53133
gttcctctaa tagtaagtta tcccacaaa tctgaatata attaagaagt tgtattogtc	53193
tgaatgttgt gtctaaaagg gttgattgat gaatgatggc tacatgtgag agtttgataa	53253
caacagctag cttagcatta gccaaagcac taactagaca ttagttttgg ttggttgtca	53313
gacaaaccgt tagacctgag aacgaaagcg tattaacaa aagatgatat gtgactttt	53373
aataaaaaa gagatggaga aaccaaattg agatttgata ggtgaactat aaatcatgac	53433
agtgcattag acaagttggt agagtttgtt actaactcat cagatttctta agaaaggcaa	53493
aatagaaac tacaccacat gtgcgtagcg ataacgtgca atttataaat aaataatggc	53553
ttcattttca tgggttagtta taaattaatg ggtcacaatt cttaatttat taggaacgta	53613

tacttcattt	tgagagtgtg	taaagttgga	agaagaaaag	ggatatagaa	agaataaaaa	53673
aatggattta	tctaattcat	cgtaaatgaa	aatgagatta	aatcattcaa	tcttcattga	53733
ataatagaat	ttaaaaaatt	gtcttattct	gaattgtatc	attaatatata	taactatcat	53793
atttaagtga	ttatctttct	tatcatttat	gtataaaaatt	aaaaatttat	aattaaaaatt	53853
atattaaaa	acataaatat	atgtaataga	attataaaaa	ttaaaattac	agctatatat	53913
aatttctcgt	cacttagact	tgacagtagac	aattgtttgt	agttaaaaaa	tatgaactgt	53973
agtcgggcta	cgactattaa	tttcaagact	tatttaacag	ttataataaa	ttttttgtca	54033
tctcgctccc	atttgtcttg	cgtgttaata	tatgttataa	atagacattg	atatattttt	54093
ttatgtgtat	ttgttgatac	ataacaacaa	taggattgta	attcagattt	taatatttta	54153
ttacatgttt	atgtatttta	ctataatttt	aattaacaat	tataaagttt	taattttata	54213
cagagaaatt	aaaaatgata	cattaaatgt	aataatcata	atattaatgg	acagaattca	54273
taattagaca	aatgtttaaa	ttgtattggt	tattgaagat	gagatttttt	attttcattt	54333
ttacttgggg	tcttagacaa	acctaataaa	aagagaaaat	aacacatgta	ataatttaagt	54393
gaagtaata	taattggaaa	agatgacgaa	attaagatga	gaagaaaatg	ctatgagatt	54453
gaaaagatat	taagtgtctc	tttatataaa	actcaattaa	ttacgagtta	actggtcacg	54513
catgaaagtg	taaaacattt	ttatgttacc	atttaattat	aaattatagt	taataataat	54573
tttaaaataa	ttatcataaa	aattaataaa	tttactgtac	aagatgaatt	tttaattaaat	54633
agtgatataa	ttttttttta	cactaacaag	tgtataatcc	tttatctctt	taattattat	54693
ctaacttttg	attccacggc	atgtcaatat	tttctctctg	accaaataaa	ctatcaagggt	54753
tagtaaacga	atataaagac	aatcccatca	tggtcttttg	tgtcaaaaatg	aggccttctt	54813
aaagatcacg	ctcaatgata	tgtagttttc	taagtgcgta	aaatgcattg	taccctcatg	54873
aagctataat	agggttcaga	gatagcttta	gaagttcaat	agagcatgtg	gacctggggag	54933
tgaggtcgta	tgctgctata	atgctgtata	aacttttggt	gagcatgcac	gaccttttta	54993
ctaactgggg	ttccatagtg	ggtttcagtg	atagtcttca	taagtctcatt	agtcctttaca	55053
agttcaatag	agagaatata	tggggacctc	ggagtgagga	ggttgaaagt	cactataatg	55113
ctgtataaac	tttgggtgtg	atatgcacca	tctgatggcc	atccaatgtc	ccctaggggac	55173
aacagggtac	ctaattaatt	ggtaccacaa	cggggagaaa	atcaacacgt	tgttgggaata	55233
tacataccta	gaattgaagg	gctagctcaa	tcaagctaaa	cttgaattca	actatagaaa	55293
ttaaattaaa	ttgaaatttg	gttacacgag	tcaggaccat	tagttataat	taaaatgcgt	55353

tagcacaatt	tctaacgcta	taggcataga	agcactaatg	gtgacacaca	ctagtataaa	55413
aatactttta	atatcagtta	tttttagattt	ttttgtttgt	gtaagtcaat	caatttttaa	55473
agttactttc	aatcaactt	taacaaaaac	taatgtagaa	atgatctaga	aaactttttt	55533
tttaagtctt	aactctttt	catcaatgtt	atcacatata	atatatatat	atatatatat	55593
cccaaaaata	ccaatcaaat	aaactactta	gtttacttat	atgttaaata	atcgacctat	55653
tacaagagag	gaggcttaac	tcagatgatt	tattataatg	ttaagttatt	ttaaatcttt	55713
taatatTTTT	atttgatttc	tatgaataaa	aaaaaatcaa	cttaccacca	tttaaagtct	55773
caagtcattg	gatattatct	tttttaata	tatatatata	tatatatata	tatatatata	55833
tatatatata	tatatatata	tatatatata	tataaattga	acatttgttg	ttatattagc	55893
taatatgtga	cattttattt	tatcttaaat	aacatattat	ctatcaaaat	aaaatgttat	55953
aacaataaat	aattcatttt	tatttgtatg	aattttcata	aattttatta	ttttcaaaaa	56013
ctttgaactc	aatgatgtga	tgaataaatt	tatatcttta	attaatgaat	tgatcgtatc	56073
ctatatatgc	atagtatatg	aaaaatcaat	tctcttaaag	taacataaag	aggccctttc	56133
gtttaagtaa	ttgaattact	tagacttcaa	aaaaacaatt	gaggggaatc	aaacaattaa	56193
cagataatta	attctagcaa	actatattct	cctatatatt	gtcaaataat	tgaataataa	56253
aactgtaaat	tatgataatg	atatttatgt	tgaacacgtg	gcagtatggt	gaccaaatgt	56313
tcgtcaaaaa	tcaaaatgat	gtgtgcctgg	tacagtcaag	tacgtatctt	tgagtatgac	56373
attcttgtgc	atgtaattct	ttgctgatct	tatccatgtg	taaaataaaa	ttagtttatgt	56433
tggtatgacg	gccacacagt	ttaatacttg	agaataattt	ttcataagta	tctctaacta	56493
aactcttgtt	agatgcaatc	aaatcacttt	agttacttat	agtactgtta	ttttatttgc	56553
tcacgcatac	tgaagtcta	cggtacttat	agtactgtta	ttttatttagc	ctaattatcc	56613
atgtataaga	atcattgaat	aaatagttag	ttttaccaga	tagaaaaata	aagagggtaa	56673
ggaacacca	acctatcatg	agagctaaag	cttcacaaca	agcaacgaac	agcttttaac	56733
cttaaaactg	gctaagtcca	atattaaaga	agaataaatt	aaaattgtaa	ggctgggtgt	56793
gtataaatta	aacaaaaggc	cctctattca	aaccttcata	tatcacacct	gttttttaatt	56853
aacgcggact	actttttcat	ataaaaaaaa	gatcattaga	ggattaattt	aaagcgtttt	56913
agtttttaat	taccaaagag	aataattatt	attaggcgct	ttgtccacca	atcaatcacc	56973
taaacagaa	aaagaaaaag	aaaaaaaaag	tcaaatgga	ctaagtcaaa	agtggcacaa	57033
tctttgtctt	gaactcttta	attagcaaca	aattatactc	ttctgcacaa	atcacagaa	57093

taccttacat gaaaagaatg ttaatttgac ggtttacatt aaattatatg cagttttctg	57153
caggtaatta attttcaaga atttaagggt gttggtaatt ttcaatagct agcttgacta	57213
gcaaaggaaa gaataaaggt aaaatgcttc ttgttttggc ctttttggat tgttatactt	57273
tttgctaaac ggaatgggtt atatgaatgg taaaggagat aaattgttac atagtctaaa	57333
attgttatag tcttaataca tcttcaaaca atatataata taattttatt aactattttt	57393
tataaattaa aatttttcaat tatatgtcac ataaaaagga attatgacca taaatatata	57453
atagtcctaa atcatgtaat aatttgtcag caaataaga gattaaataa agttagaagg	57513
aacacaaatca gtggattaat taaaaccgt catataaact aaattaata tcaaatctca	57573
aagtacgcac aaaaaatctc aatagatttt tgtaataat taactagtat tcttaaaaaa	57633
ttgattaaaa aattaaaaga aaaatatat ttattatata aatcagaaaa atataaaaaa	57693
attataataa atacaatttt tttcctccta ataacaatat ttttatttga tgccttaga	57753
atatgggct aacatttaac atataattat acttgggaaa aatttactag aaatagtaat	57813
aatagattct ctaacacttt ctccaaacat agtctatgat taatttaaat ttattgaaaa	57873
ctatgaagtt atgagagaat tattattttt gtaattttta agaaatttca actagagaaa	57933
atatgtttaa aagagcctgc tgttaaaact tctcaaaatt tattttcaac ctctaacgc	57993
agacttctga aataagcatt catgcacttt attaactagc gggtgcaaca actcctatgg	58053
gtttggaacc aagttaagtt tccctttggg gtctgaacct aactaaatta acctaactg	58113
cctaacctca aaggacttta tcttccccc acctctaact caccctataa aagcaccctc	58173
tccactctt acttgcattg caaccttaac ctccagcatt caccactaagg tgttctctgc	58233
tcgccaaaag atcatggagc ctgccaaaac cattcacaac aatgtcaaat actccccat	58293
cttcttagcc atctttgttc tgatcttagc ttcagcattg tcttcagcaa atgccaaat	58353
tcacgagcac gagtttgtt tacattctct cactctcttt ctcttaattt ctctggactt	58413
attttattct tgtttttttt aactcttttc cgttagataa taattaccta gctgttgatt	58473
gtaattgaaat aggttgaagc aactccagtg aagaggctgt gcaaaacca caacagcate	58533
acogtgaatg gacaataccc gggcccaacg ttggaaatca acaatggaga caetttggtc	58593
gtcaaaagtca ctaacaagc tcgttacaat gtgaccattc attggtataa tatcaagcta	58653
gcactttaac ttcatttttg attatgaaga ccttttgact taattttaca catctgttta	58713
ccacgattta tagacatata gatattagtc ttacgaaaaa catttttctc actttataat	58773
tatactactc ccttactctc tttttatatt aaaatggaaa gagtctctct ctaccatgtg	58833



aaaaataata tataataaa aacatatata ctctaccatg tgaaaaataa tatataaata	58893
aaaacatata tacacttttt acttcttgac attctaattt tttttacctt tctttctttc	58953
ccatgatatt tatcatttat tctctcctct attctcattt atctatctcc caaggtgtgt	59013
acattccatt agaagtgtgaa aatgaaaaac attcacaagc ataatgtaaa aaaaataata	59073
ttattttctca taaccttata tatatatata cgccacataa tacgtacgaa cgtaagtgtat	59133
actatcatga aagtctttga atggctttct tttcagggtg aatacatata ttaatggata	59193
gtggtttttg ttggtcattg tttcttatta ttatgtcctt aggcacgggtg ttaggcaaat	59253
gagaacaggg tgggcagatg gaccagaatt tgtgactcag tgcccgaatt gtccaggagg	59313
aagttacacc taccgtttta ccgttcaagg acaagaaggc acactttggt ggcacgtca	59373
tagctcatgg ttaagggtca ccgtttacgg tgctttaatc attcgtccta gggaaggaga	59433
accctaccct ttccccaaag ctaagcacga gacaccattt cttcttggta ttttaatttc	59493
cttcttaatt tactcatgca tatgcattat ttgtaattat agcattcatg gtaacatgga	59553
ggcaactatc taaaaaagaa ttgagattat ttaaaaaacta ataagtgtatt gtgatgttg	59613
tgattaatta attaatacta ttgaagcaaa gagacaatat atatagaaat tgtggttttc	59673
tgttgtttaa ttttctttt ggacaaagat taaacggtta aagtgtgatg ggtgatgatt	59733
taggggaatg gtgggacgca aacctattg atgttgtgag gcaggccaca cgaactgggg	59793
gagccccaaa cgtgtctgat gcatacacta tcaatgggtc acctggtgat ctttacaagt	59853
gctccagcaa aggttttgatt aattgcttct taatttcgat tgcattaatt gaacatgtca	59913
catgtcttgt ttaataaat ttactttgca aaatatttga cataataaa acagggatag	59973
cagtcataaa aaagagaaaa cgacatatga tatgaaatta ttaaagatgt caattattta	60033
tgaacaagt caacaatagg ttgccttttg gtgcagcgtc tatttcattgg cctttctact	60093
ttttgtctc ttttgaagac aaaagtgtct tccccaccaa ataaaaagaa aaaaaaatgc	60153
agaagacttt taagtaata tatagtttat aaattgcaag ttttagcaag aatttttaaa	60213
aatataatt tgatatTTTT tttcataaat tagaagaaa aggctaacac tttttcttaa	60273
aaataattat attataaaat tgtctattta aaaaaaaac cattgaaatg acttaaaagcg	60333
aagagatttt atacgcgaaa cctgctttta atgattttat gcaaccaaca aggtgtcctg	60393
caagtcaaat ggaaaaaagg cattttaaaa cataaagtta atcaaacttt tcatttcttt	60453
aatttagatg atgtatcatt ttaatttctt acattttctt aaaatattaa ttttatgcat	60513
tttcaacata acttttttta tatattcaac taatgagaaa atatgataaa tataatttta	60573

aaataattat tgtaaaaaa tttttttact ataatatta atttgtaaaa aaatctttac 60633  
 actataaatt atttactatt cttttttttt tacttcacat ggattccttc ttaaaacttt 60693  
 catttttttt actcaaaatc tgtaaaaaa tccaataaat tgttttaccc ttttggttcc 60753  
 atgcagacac caccattgtc ccaatacatg cggcgagac caaccttctt cgtgtcatca 60813  
 atgctgcact caatcaacct ctcttttcca cgtctgcaaa ccacaaactc acagtgggtg 60873  
 gtgccgacgc ctctcacttc aaaccttcca ccaccaagt cctcatgctg ggccccgggc 60933  
 aaaccaccga cgtcttaate accggcgacc agccaccttc cgtctactac atggcgcgcc 60993  
 gtgcgtacca atccgccccaa aacgctgcct togacaacac caccacaacc gccatactcg 61053  
 aatacaaatc accgaatcac cacaataagc atttccacca tcatgccaaa ggagtaaaga 61113  
 acaaaaccaa acctataatg cctccactcc ctgcttacia cgacacaaac gcagtcactg 61173  
 ccttcagcaa aagcttcaga agccctagaa aagttgaagt acccactgaa attgaccaga 61233  
 gcctcttctt cactgtgggt ttaggtatca agaagtgcc caaaacttc ggaccaaaga 61293  
 ggtgtcaggt attggactat tcacctaat ctattatcat gcatcaattt aatttgcatt 61353  
 tacgtatctt atcttaagat ttcaataaat gtctcatata ggaaaaatta cttatttatg 61413  
 tttataatcc ccacaaattt tacattttta tcatactct taaaatttaa gtctaattta 61473  
 atttcttatt ctttaaaat gactgatag ttctgatacc aaagaattca aatattaaat 61533  
 atttttattt tttgtcttgg tattctattt tttcataaat tctaactctt ctaataattt 61593  
 caattcatat taagatcggt aaatagaaaa tctagaaaaa aaacaaaaa aggtattttt 61653  
 ttttcaattg ttttattttc aattgatttg tcattaacaa actgattctt cttaaatctc 61713  
 acaaaagtac atgtcgatat aaatatgaga ttataaatcc atgatatcta ttttcgattt 61773  
 ttacatataa tgtttttttt atctttttta gttcctaata agcattttta aatgtcttat 61833  
 gttctacttt tgcatactcg ggaccocatta atgggacgag gttcactgag agcatgaaca 61893  
 acgtgtcttt cgttctcccg aacaacgtgt ccatcttgca ggctcaccac ctcggaatcc 61953  
 ctggagtgtt caccactgat tttccgggga agccgcgggt gaagtttgat tacaccggca 62013  
 atgtgagccg ttcgctgtgg caacctgttc ccgggacaaa ggacacaaag ttgaagtttg 62073  
 ggtcgaggtt gcagattgtg ttgcaggata ctgacattgt cactcctgag aaccacccta 62133  
 tccatcttca tgggtacgat ttctacattg ttgcagaggg tttcggaac ttcgacccaa 62193  
 agaaagatac ggcgaaatc aacctgttg atccaccttc gagaaacaca gtggctgtgc 62253  
 ctgtaaatgg atgggcagtt attcgatttg tggctgataa ccagggtaaa taaatagggt 62313

cttgtaaatg ggcgtaacat tagttaggaa actaaatata aaaaatattt attgtatatg	62373
atataagaga atgtaaaaaa atttataaaa aacaactttt catattttta taaaaaaat	62433
cttttaattt cttaatcatg ttctaagaac acaaaattaat atttaactag taatatattca	62493
ctattgagtt ttaattagat aacactctag aaatattatt tatttatata tgaatatata	62553
ctttctttcc agttttcatt tgtataatgt cttttttata aaaagaaaga aagaagcaaa	62613
ctcaatcatt tgatgggtgt gtaaatgttt aattgtagggt gcatggcttt tgcattgtca	62673
cttggacgtt cacattggat ggggtttggc tacgggtgtt ttgggtggaga atggagttgg	62733
gaagttgcaa tccatagagc ctccctctgt ggatcttctt ctttgttagg atatcatttc	62793
aaaatatctg ttggccccc acaatcggag ttttgcagtt tttcttagtt tggaaagctgg	62853
ttgatgcttc cctgcattaa ttttgggagg gttttttgtt ttgcttcatt gctttgtttt	62913
aattgtacgt tctttttcta gagagagata aattgggttt ggaacctgag agagggtgg	62973
atgatcggtt ttcaacgcaa ctataccatc tgaacagtc ttgcccaact agttaaatg	63033
atgttatttt ccccccatag ctcaagttac aaacagattg caagttttga atatcaatat	63093
atcagctggc ttttttttgg gttgcaactc ttgggtctct gaataataat atgttgacat	63153
gtattaagtt tcaaagtcct ctaaccattt gccatcataa actaatcaa aattctcaca	63213
ttaattcaca gtatagtgt tacagagga ttgcttatcat agcatcaagc attaaatcc	63273
ctcaacgaat gttgtcaaca tatggattat gggataaaaa tataactacg cttttctgat	63333
agaggggtgg ttttggtttt catcaggctg aatgtgcaag ctgatagagg agtattaaga	63393
tagagttagc tattgtaatt gagaagcaga aagctactgt cgtgatcat gtgagtgagg	63453
ggcaaatgta ctagtccagt agtcctcttg agtttttagc caataatctc atcaaatgaa	63513
tctctgggct tgcttcatta atgctaaaag ttgctctctt ggcttatgct tttttttct	63573
ttgttaataa agtgatgaaa ataaaggata tattcataaa cttttattat gtttactttg	63633
tataaatttc ttaacaataa cttattgaga aaaaaaaatc aaacttctct cataacataa	63693
aattaacttg tgcatttcta actttttaga agttttctta ttaacttcc ctaaaaaaca	63753
ttacatttgc tacagtaact tgggtgacaga tttcatgtta gctctcgcat gatttcaatt	63813
gtattacaac attttgactc tcagttatcc ttaataaaat gatatgattt tctttgccta	63873
aaaaaaactt ggtggcagat tttgttgcat agattaaatc atctatgtgc ctttaggcat	63933
ttgctagtca tgatggtcta taaatgcagt ctttacctgt atattcttag attttaacat	63993
tgttgggttc tgtactatct tttggggtct ttatttggat tcaccttggt cttaaaccta	64053

aaggctattg atcattttat tcttttctt tttgaatttt ttccattaaa gtttttctt 64113  
tttataagtt gatttttaatt tttagaagaa acttaattca tttctctctt tcttttctt 64173  
cctgatgaaa aattattcaa attgactttt aaaagattg tatcatatat atgtgacatt 64233  
ttctttttca ttatttttta tgttcatgtc acataaaatg caagtctcat aaaatgtggt 64293  
tgtaggacaa aaagtgtac tatactccac atggtatata taccaaaata aaagtaatac 64353  
acggaatcat gtgaagccta ctcaagtaga tgtgcaaaat cttgtgaatt caaattagtt 64413  
gtcttgttta ttcattacct tttcaatttt tttaatcatc ataattaagg ccttctgaat 64473  
ccctttaagt gataaaagaa acgtgcaatt atgcaacaaa taaattttcg ttatgttact 64533  
atthagtcaa ggaggaaaa aaagtgtata gggaagaaac aaggatatt tcctgttata 64593  
acaaacttaa aatggcgact attttgacga cattgcaaat actcatagta cgatataaat 64653  
tttgaattta atatacaatg aataggcata ttcattttct acccaaaaa agcactactca 64713  
tttatgtaca tttaattttc tctccataga ggaattaatg tacaacatg cataagggat 64773  
gagcaaaagg gacagattat tgcaatccag aagcatccaa ggaaagttgg ataaacaaat 64833  
caattaatat atataaaaa aaacaaaaa tgcctcctagt aaaagattaa aggaagagtt 64893  
ggctatatat ggcaaacctt ttctaactgt ttaccctct tctcatcacc gcattgcac 64953  
accaatcagg gaactttccc attacaaaac tcattggaag ccaacatate ccccaaaatt 65013  
cccaactgat ctgcattgtc catgaaattt gacattttct ctctacaaa attccccatg 65073  
ctatgtcgtt ttcccaccat cactaggcca tagtcctctt ccattccccg aatcgcttct 65133  
aacacttgta tgcaatcttc caccacagcc tcattggttaa caacgttaac actatcacta 65193  
ctaataatcat tcttggcaat aaactcatct attagactct cgtccaactg gctctctaaa 65253  
ttctcatctt cattttctac aagtccatta aacctcgaat cttcattggt aggcaaaaca 65313  
aaccggaaca aggtgacacg tgtgttgggg cgttccaaca ttccaatccc caacgccaat 65373  
gcttctctat catcttttcc accaatgaag aagataccaa catcaaaaga caacttggag 65433  
ctactccac tttagacoga gtatctatcc actagtatcc ccaaagtacc ctttgcatta 65493  
gcgagaaaaa ttgtgttgag gttcggaatg gtgctggcca ggtggcttcc tagggtttga 65553  
tcatttttgt gaaagggtat aatgagaaga tgcaactgag tatcttcggc taggttaca 65613  
actgcctcat gcactcttct gtacggagcc acgttaacat acgaaagaac cgttacagga 65673  
ccacttgagt tgttggagta gttctcaagg gcacgtaaaa tgtggttggt gttggggtag 65733  
tttacggaca aagattttcg cttgtttttg ttcattggga gaagaatggg tgtgcttttt 65793

ccaacgagct caatgagatg gaccacgtag acgtacaagg ggctttgtgt ggtagggttg 65853  
 catgcttcta ttagggaat catgttgtgc acgtgttcgt ctgtgtgcac acatgatacg 65913  
 atgttgaaoc gtgtattttc ggttatgttt tggattgttt tcacacaccc ttcttgtatg 65973  
 gtttgctct tacacacct acggtgcctg tacaaggatt tgatcaaggg tatgcaaatt 66033  
 gatgtcatga ccaccacaga catcaccgct acatgaata ctctgtatc aataacctgc 66093  
 aaaaggaac gtcatgacat aaattagatt taatatttca ctaacaatat gtatagcacc 66153  
 cacattaaaa agcatgagag tattgaaatt caaatcgctt caaagaatt atcatatgac 66213  
 ccgaactat catgtattct tgattccgat tgtcaatttg tattacatg catataatca 66273  
 attttattat tttttatat atgaaaatt tgattatata tatcttgacc gaaatcgtag 66333  
 acatgtagt atgtcaaact ataaaaaat tccatcaatt aaaacttac atgcataatt 66393  
 aatatTTTT aattgagaga cagaaattga gaaattatta gacaatatga catcatcaca 66453  
 tgattatgat ctatttcagt ttctacatga tatataatca tgtttttttt tcttaattta 66513  
 caaaagaaag aataatatat atgactttat tatgtgtcat atagatgtta ataagaatcg 66573  
 tgaaaatttt agaccacttg atcatttctc gaggaagag catgaaaagg agacaagcga 66633  
 ttgtatttct acctctcata aaatagaaga aaatatataa gagaatatgc attacctga 66693  
 gtttattcat tctaccgtaa aagatgagtt cgactatacc ctgacattc aatatgagcc 66753  
 caagcagac gccatgtttg ggcttaatat tatagggtgg agaaatgagc gcacatgcaa 66813  
 gtaccttcac caagcatccc acaacaaga tagccagaac tactagaacc acctccaat 66873  
 gtcatgaat taaggatcaag tctgtccttg tgccaatcaa caagaagaag aaaggcataa 66933  
 aaaactcata cagatcaat tcacttctct ctatgattgt tgtggcgagg ggggttccat 66993  
 ttggcagaac caaaccatag agaaaaggtc ccatgacgaa atatatgccg aatgtgtcac 67053  
 taatagctgc cataaccaa ggccccagaa gtaacaagac aacgtatgct tctttaattg 67113  
 gcttccagg ttgtgttctc tccaaaacaa tgttactag ttgtgtata ataagaagta 67173  
 ataaacagc aaatccagtt gcacctatca aaagcacgat tagaaatgc atgtgaatt 67233  
 ttgagtgaa caggagtcc atcgtagtcc attgcaatat ttcactgatc atggctgaag 67293  
 atagagcaat ttggccaagt tctgtggcta caagggttag ttccatcaag gtttcggata 67353  
 caacagcga actgctcaag gtgaagatgt tcgggaagtg gtgattgac atttggtttt 67413  
 gattagcatt accgttggga gaatacaaa aaaacaaggt tacagtgacc aagaagaag 67473  
 ccaagaagg aaacacaccg aatcgccaac aacgtttcgc tgattttaat gttgtacca 67533

cgtccatctt caaacaagtg aggaacacgc aatatgtagt gectatcttg gacaacgtat 67593  
 ttagaacacg ggaattgtctt accggaaaga gggctcccaa tatttcttca tggcgcccta 67653  
 agaattgttg ccccaacaga atgccagcct gtaaaattta aacgaaatta agattattta 67713  
 tgggtatggt ggatagaaaa aacaagacca ataaatataa aaatgattat atatgtgtat 67773  
 gtttctcttt tttgtgggaa aagggaacgat aaacttctaa aaaaacctta agaaaacacg 67833  
 cttgaagaag cgtctgccat aatgcogtta aacaagatag aatgatcaaa ctgtgtcagc 67893  
 tcgtattgaa cactagtaat taaaattact aataaattaa cctatatatg agaaatgtag 67953  
 aatatattag gagtgaatgc ttacgatgac gcagcagatg aatttgggtg tattgatagg 68013  
 cctgaggaca taatgaaggg ctttagagag taaactgact aaaatgattt ggaacaatgt 68073  
 gacaggaacc acaaaaatga atggattatc acccatccac acgcctaaag agcctacagt 68133  
 tctatcatte ttgagacaaa ctatcatttg acccgttcta gcgtcattaa agaccgtgggt 68193  
 aatggaactt tcactcatta tatatatggt tgtaccaagc aattagtgtt gacaagaaaa 68253  
 acaaaaacct ttctctcttg ggtgttggt gtttgcaaca aacgcaaca cgagcactac 68313  
 ctgaggggaga acaaaagtag ataggagaaa agcatatatg tgcattttta tatgctctaa 68373  
 tcaaatgtgc agataaaata attatattta acaaaatag agctgcgaat gtggtgttag 68433  
 ttattacaac ggttcaaaaa tcaattatag caatgataac tttagtcaaa tgtattccga 68493  
 gttggagttt cgagaaaattt tggctaggaa agaaaattga tatgctatca ctgagctaag 68553  
 gccagtgtgt aacgaaaaac caagaagagt gaatttcaat ttagtctctt taaggataaa 68613  
 ttttaaccaa ccgtgtcaca aatatcgcac aacaaaagtt tgtgtacaaa tgattttaatg 68673  
 aagacagggt tagtcaacga gaataactgg ttacaaccaa caatgacata cggaatagtg 68733  
 gcagggatga tcggaaaaaa aatacagcaa ataggcaata gtaaaattta atttgtgcgc 68793  
 cgtttcacga taatattttc ttgtagaaaa ggattctcct tgcaatagag gagtacaaaa 68853  
 agatttgtac aaactactac tagcatttaa taataaaaaa taattttgat aacaaatat 68913  
 aaattaaaa atgaaaattc atgcatttaa aattttaagt gttgaaattt ccttaaaaaa 68973  
 tagtttgatt gattgatata atatttcttt attttgatat atatatatat atatatatat 69033  
 atataatacg ataattgata gtaaaatttt acttttactt ttaaatacat atataaaat 69093  
 taaggaaata gtatgtgatt aattttacta cttttctata atatagttga aaataagaaa 69153  
 taaaatcaag acacgtactt gtttgctatt attatttaaa ttagaaaaatg aaataaatat 69213  
 tttattgaac ctaagtgctc tttaatcatt tagcatggag atcatttgga aaagtgtgta 69273

taatttaact agatgtccgt ttgaaaagt ttatgtgtat cttaaagtta tattattgta	69333
taaagtgtat gtttcgaaga acttattataa ataattata atatatatct atgataaaga	69393
ctaacaataa ctaaaaaaa ctttgcata cattgttaat cttttacatt aatttaaaat	69453
ataattcaca ttttttattt ttattttatt atgaatttta attataatc atattcaaaa	69513
ttattttatt attataaatt ttagttaaat aaaataaca tttatttttc aaattctaca	69573
agctaataaa ctagtattac taaatgcata gagtatgagc aaattctgtt aggtggattt	69633
tgatatattat atatgttttt cgttaataaa ttattactcg actttagaag gtgaaaaag	69693
gtcataaaaa tttagctatt atgaatctaa actttctttc tttttttctt ccgatctttt	69753
tgttttttca tcacattagt gatcgtatat gttatttttt ctgatttttt ttctatctc	69813
tttattctct ctacctata cactaaaaat gatatatata catttttcca caaagactcc	69873
tgcccttttc tatcagctaa aattatttat gtacaaataa aaaaggtaaa aacacaacat	69933
ttatttatga acagataaac gtttttgtga gacattaact gaacctactc tatcaagctt	69993
attattacta ctactactta tcttctctcc accacactgt gtccactaaa ccggaaccat	70053
ccccatacaa aattctactg aagacaacat atccccaat attcccaatg catcagcgtt	70113
ctccatgaaa gttgtcatatt cttttccatt caaagatcca tcattgtggc gccttccac	70173
catcaaga tcattagttt cttccaaact atgcactgct tccaacacat ccacccatc	70233
gtccaccgta atctcgtacc aacaaacgtt accaatgcca tatttcattg ctttgaactc	70293
gtcaattaac ccctcgtcca acatgggtatc ttccctctcc tcttcacgct cttctcttgt	70353
caaaataatt ttacaaccac acggtttctt gttcacgata acaaacctaa acaagctcac	70413
cctcgtatct gcacgctccg acattcgaat tcccaatgcc agagcttccc tategtgggc	70473
cccacctatg aagaatacac ccacgttgaa atacatgttg ttgttgttgg acgcgccag	70533
ccgagagtgc cgggccacga gtatcccaa cgtgcattgc gcatgcctt gaaacctagt	70593
attcatcttc ctgatggagg ctgcccctgt tccgacaagg tcaatgttgc cgttttcgtg	70653
aaaggggatg ataagtaag gcaccatatt gtcttgggag aggttgaaaa ttggcgtcgtg	70713
catgctcttg taagggtcca cgttgatgta gggaagaacc ttgactggcc cacttgagtt	70773
gttgagtag ttttcgaagg cttgcatgat gtggttggtg ttggggtaat tcacagacaa	70833
gaattttctg cgaccgtgtc tatgttttat gggaaggaga atgggtgcac ttttccccac	70893
gagctcgata aggtggactg cgtagacgca tatggggctc tcttgcactg ggttgactc	70953
ttctaataag gcagtgatgc cagcacgtt tgcttcatta tgtacacaac aaacaatgtg	71013

aaactctctg	tttcttggag	tgttttgat	cgttctcagt	tccccttoga	ataagcttct	71073
tgcgtgtact	cgagggcgat	gcttgtacaa	tatgttaacc	aaggggtgta	caatcgcggt	71133
tataagtacc	acacagaaca	ccaattgact	gaatgtatcc	tcatacaaca	tctacaaaac	71193
gtaataataa	cataaattag	atttgatttg	atgtattgta	catacaatgc	atcttcaatg	71253
ttttttttaa	tgaattagt	cttgagtta	tatttcagt	ttattttaat	ccccaaagt	71313
atggaagagt	caataaaatt	ttgatttttt	tatgatttaa	ttcgagtctc	tcaaattggc	71373
gtcttttttc	tgttttgggt	tcccagtact	taagtcataa	taataataat	aataataata	71433
ataataataa	taataaaagt	ttaggaagaa	aaatgagaaa	actaattagt	ttogggtttt	71493
atttaagggt	tttttagttt	cagaaactaa	aatgacatat	agatacaaat	ttaaagatta	71553
aattcgctat	ttactcat	tttattagac	aaaatttagg	tttaagcatg	catacctttt	71613
gtttcttgaa	tctagccaaa	cttatgagtt	gagttatacc	cttgatgttc	aacatgaggc	71673
caagtagcgt	gccatgttta	gggtctaagt	tgtacgctag	agaaaccaac	acacaggcca	71733
gcaacttggt	caagtctccc	gcgaagaaaa	cacctgaag	ggtaaaaaat	aagcgccaat	71793
cctcaagtgc	agacaaatcc	gtgttgatgc	caatgtacac	aaagaagaag	ggtagcaaga	71853
actcagtggt	aagaacttca	ctcttctccg	ctagcgttgt	ccctaagggg	gggccacttg	71913
gtacgacca	accaaaaatc	aagggtccca	tgagaaacgt	tacgcccatc	aagtctccca	71973
caccagccat	cactagcacc	ccaagaagta	tcaacacaac	ataaagttcc	ttcacagggt	72033
ttccaactgg	ggttctcata	gcaatcaatt	tcattggttg	tcgaagaaca	aagaagttga	72093
agaagacgag	caagcacca	tttccaaca	atgcgattga	ttttttcaca	tcaacattag	72153
atgtgaagct	atgcataacg	atgaaaagcc	atagtatgat	gtcgttgatc	atggagggaag	72213
agagagcgat	ctggccaagt	tctgtggctg	tgaggttcag	ttcaagcatt	gcattcagaca	72273
cgacggggaa	gtgtctcaac	gacattaagc	agcttactga	gacacgcgcg	attgttaaag	72333
atgcagagga	aatttgttgg	gggtgataat	agagacataa	gagtgctaag	ataaccacaa	72393
atgaagcgag	gaagggaatt	actccaagtc	gccacgtgct	tttggctgct	cttatgtgtca	72453
ttagtacgtc	cattttcaat	gcaactaaga	atacgaataa	tacggcgcgc	gttaaggatg	72513
ccatcactaa	atactcgtgc	tgtctcgttg	gaaataaaac	ctgccaaata	gtcttgttcc	72573
gccccaaaaa	tgtggggccc	aaaataatgc	cacctgcac	attttcacta	gacacaaact	72633
taacattcaa	actaaacct	taattataat	ttacaataca	tacaagaatg	aatttcttct	72693
gctagaatta	acagtttcaa	accttaaat	atactaatag	atgggttaaga	tttctactta	72753



aatcatata tatcatttgc tttcgttttc aatgctaaaa tgatgtcagt aataagcaac	72813
agtaataatc acattcatga taatagtgat aaaatgatca ccacagtaat atcaatcaca	72873
atagtcaaaa taataataat aatgaccttg atgtgaaaaa tgctaaagtg atttttatat	72933
aaggaaatca ttctcatata gaaatgataa aattacttat tatgagaatg aaaacaataa	72993
attcttattt gaatggttag atttaaaaaa tacatcactt cttattaagt ggatcatgtg	73053
gaacattaaa ttacctttaa tctttatcat aattactctt ttcattotta aattaagatt	73113
ttttttctaa tttctagata tattaattat ttttttctta aatattotta cttaattatt	73173
ttctcatcaa atattaatga gatgaataga gaaataagaa aagaataatt ttggaatgat	73233
aatataatta attaattaat aaatttaatg tgattaatta aattaattat ttttcttaag	73293
acacataaat tagttgaaag gtaattgtaa taaggagacg acggagtga ataaattgct	73353
actaatttca taaataaaaa accttcaaac tatgagcaat cgtggctcatt gaaaacaaca	73413
aagaaatgaa atcctaactt tttatcatga gataaatttt cttggcaaaa tgatagacag	73473
acaaaaatgg tgtcttccat gtctctcctc ctccacaaa caaacgttac cccatacctc	73533
agacactcct cgacaactat tggacttgtg gagtccaaca ttggactcct tgaagaataa	73593
ttcggaacag cctcaactac ctccaatttg gaactataaa ccacgttata acactgcctt	73653
gtcggagaag tgcacctatc cagcttatgg cgctttcggc tgttttccag cggcgagact	73713
aacacattaa tagttatgta tttcaagggc aaaagggtcca aactctagcg tgtggaagg	73773
cgtgaggtgg attccatggg agaagaagtc tcgcatcatg ttgaggaata aatggtgtat	73833
aaggagaagg ctaagcgtca tgtttgagga gactaatttc ttgaagcaa tttaagaaa	73893
atcatctttg aatatattgat ttttaagaca gttttaagaa aatcatcttt aaacactcaa	73953
tttttaagat gatttttata aaattgtcat catatcctt ctattattta caaaattatt	74013
actgcctaac attttaagac aaatttttaa aattatctta aaaaatacgt tgaataaatt	74073
atttttttagt agtataatca tatttgtcac tgttttccatt atcattatta tcaatatttt	74133
tattgttata actaccatct tcacctatca ttaacatcat cattccatcg ttgtcaacgt	74193
gagaagggtg tagcttgata ccgacatgat gacaatgat acgatttgtg ataattgttg	74253
caaaaaaagt gaaaagaaaa taagggttta aagagagaga gaaaattgta gcctttcaat	74313
atttttttaa atttattaaa agggaaaaaa acttttatag ctgacttacc aggacgttgc	74373
agatgaattt ggggtgtctt agaggcctga gaagaaagta aagcgttcga gagactaaga	74433
tgacggtgaa gagttggcac aacgttacag gaagcacaaa ttcaaatgga cgatccccaa	74493

tgaaaaattcc taaagagccc acattttctat cgtcttcgac acaaaactgc cactgcccac	74553
gggaatccca ataacttgag acaattccat tgccctgtga agtcgccata ttatctctat	74613
atataatccaa ctaatcaaat tagtatttct gtgtttgtgt ttgcatgcat gatgacgcaa	74673
cacaaacacg taccaggga cataaaatgc tgagtggagc acgcaatgat atgaaaaaac	74733
aatgtcttaa tttttgttg ttgagttatg aagaacaatg tctgaatgaa acaaatgaaa	74793
aaagaggagg ggaggggtct tagaagcttt agattaggaa tcggaatatt cgtcaacgcc	74853
aagatataat agccttgcat tgcattgtga tattttcgtc aaagcatgag agttaattgt	74913
gttacggggg cactaatcta atattatagc cactcgtgta actttaaaaa aatttctatt	74973
ttggcaaaa ataataatcta tatatggatt gagaaattct acttaaaact agcgtaccac	75033
taaaattaat ctcatatcgaa gtcaacactt gaaagttaa agtaactaat taaaggttgc	75093
cctaacaate atagcccata agccttagt gggccttgtg gtgtgggcca ttttttaaaa	75153
agtattgtag ggacaaaaat aaccttaatg aaagatggga gtgaatgact ttacaccct	75213
tgctcatctc catgttttct catagcttc aatgggtgtg gacttttgtg gtaattctat	75273
ttgggtgttg tcagggaata tggaggcaag tgttcttctt ggaaacagtt atgggggtgt	75333
agctcgcgtg gctgtggcgg gtcttcgttg tggagtgatg ggctcgcca tatctcgttg	75393
ggctaaagaa caaaaaaata aataagctc tgatgggttg gaaaggtgtc aaatacactc	75453
agcatattta ataattatt attaaaaaaa ataccaacaa agttgcaagt aacatatcct	75513
aactcttaat taataaaatg cacaatctaa cagtttatat ttcctttatc actatgaact	75573
atctaacaag attcaccatc ttaaaactca tcatttgcac gaacaaaagt atttaattta	75633
ttttatctca taaatttcta tgaacctatt tttggagatg ttatatggtg aacaaaaaat	75693
attactaatt cactaagtt atgtgcaatt aatttttagt ggagtcaatt tatctttaaa	75753
atgagtcaaa ggacctattc tatggcaatc aaccttttc attagcttct attgttctaa	75813
tatacaaat ttaaatccta gtttttggtc tcaaatgttt ttaattaaat ttgttaaat	75873
attttttatt taaccttact aatctaaca tcatggttat atataaatgc aagcatattt	75933
attaagataa ttttttaaaa aattatata aatgaaatg atataagatg tatatgaata	75993
tataattatt taaatatata tatcagcatt attaatTTTT tatattttta atattaatat	76053
aaatatattt ttataaaaat aaaaaaaat tacaaatcgt aataatattt ttgtgaagaa	76113
ataataaag ataataaat ttacaaatat tataattttta aaaaattgca taaattattt	76173
aatacatgtt ataattttt aaattattcat cactgtttaca aatattttca tcatgtatac	76233